II

Challenges in Macro-Finance Modeling

Don H. Kim

This article discusses various challenges in the specification and implementation of "macro-finance" models in which macroeconomic variables and term structure variables are modeled together in a no-arbitrage framework. The author classifies macro-finance models into pure latent-factor models ("internal basis models") and models that have observed macroeconomic variables as state variables ("external basis models") and examines the underlying assumptions behind these models. Particular attention is paid to the issue of unspanned short-run fluctuations in macroeconomic variables and their potentially adverse effect on the specification of external basis models. The author also discusses the challenge of addressing features such as structural breaks and time-varying inflation uncertainty. Empirical difficulties in the estimation and evaluation of macro-finance models are also discussed in detail. (JEL E43, E44, G12)

Federal Reserve Bank of St. Louis *Review*, September/October 2009, *91*(5, Part 2), pp. 519-44.

In recent years there has been much interest in developing "macro-finance models," in which yields on nominal bonds are jointly modeled with one or more macroeconomic variables within a no-arbitrage framework. Academic researchers and policymakers alike have long recognized the need to go beyond "nominal yields only" no-arbitrage models (i.e., to include a description of the macroeconomy or other asset prices). Campbell, Lo, and MacKinlay (1996), for example, have emphasized that "as the term structure literature moves forward, it will be important to integrate it with the rest of the asset pricing literature." Policymakers have often used traditional theories such as the expectations hypothesis and the Fisher hypothesis to extract an approximate measure of market expectations of interest rates and macroeconomic variables such as inflation, but they are also aware that risk premia and other factors might complicate the interpretation of the information in the yield curve; policymakers therefore would welcome any progress in term structure modeling that would facilitate greater understanding of the messages in the yield curve.[1]

Despite much exciting work in macro-finance modeling of late,[2] as a central bank economist who monitors markets regularly, I have found it difficult to bring the current generation of models to bear on the practical analysis of bond market developments or to implement the models in real time to obtain a reliable measure of the market's expectation of key variables such as inflation.[3]

1 See, for example, Bernanke (2004a).

2 Examples include Ang and Piazzesi (2003), Hördahl, Tristani, and Vestin (2006), Rudebusch and Wu (2003), and Ang, Bekaert, and Wei (2007 and 2008).

3 I emphasize that I speak as one of many central bank economists and that my views as stated in this paper do not necessarily represent the general consensus among economists at central banks.

Don H. Kim is an assistant professor of finance at the Yonsei University School of Business. This article is a revised and abridged version of a paper that appeared in the Bank for International Settlements Working Paper Series and the Board of Governors of the Federal Reserve System Finance and Economics Discussion Series. The author thanks Claudio Borio, Stefania D'Amico, Peter Hördahl, Deb Lindner, Athanasios Orphanides, Frank Packer, Min Wei, Jonathan Wright, and the seminar participants at the Bank of England and the BIS for helpful discussions.

The academic literature provides little evidence in this regard (either for or against macro-finance models). One exception is the recent paper of Ang, Bekaert, and Wei (2007, ABW), who performed an extensive investigation of the out-of-sample inflation forecasting performance of various models and survey forecasts. The authors found that the no-arbitrage models they used perform *worse* than not only survey forecasts but also other types of models.[4]

It thus seems useful to review and discuss various challenges in the specification and implementation of macro-finance models that might help shed light on the lack of documented practicality of macro-finance models in general and on the findings of ABW in particular. To this end, I take a closer look at the role of the no-arbitrage principle in macro-finance models and reconsider the assumptions often made in this literature. The no-arbitrage principle itself is clearly a reasonable assumption, but the models also make *additional* assumptions whose validity may not have been discussed thoroughly in the existing literature. I also discuss "more advanced" issues (such as structural breaks and time-varying volatility) that require going beyond the standard affine-Gaussian framework of most macro-finance models and the challenges encountered in this regard. Much of the challenge in macro-finance modeling is *empirical*; hence, I also discuss at length the difficulties in the implementation stage (estimation and evaluation of models). Although the main focus of this article is the extraction of information from the yield curve (particularly inflation expectations), much of the discussion may be relevant for macro-finance models developed to address other issues, as they share some of the key assumptions discussed in this article.

The state variables in the reduced-form no-arbitrage model framework (on which most macro-finance models are based) can be heuristically viewed as forming a *basis* onto which to project information in yields and other data. Herein I make a distinction between models that use (what I shall call) an "internal basis" versus those that use an "external basis." By an internal basis, I refer to a basis that is determined inside the estimation; hence, it is unknown before the estimation. Latent-factor models that describe inflation expectations and term structure jointly (e.g., Sangvinatsos and Wachter, 2005, and D'Amico, Kim, and Wei, 2008) are examples of internal basis models. By an external basis, I mean a basis that is a priori fixed completely or partially, as when a specific macroeconomic variable (such as inflation) is taken as a state variable. Note that no-arbitrage guarantees the existence of *some* pricing kernel, but it does not mean that the pricing kernel can be represented well by a priori selected variables. I shall argue that external basis models involve strong assumptions, and I discuss potential problems that may occur with their use. All is not well with internal basis models either: The weaker assumptions of these models may come at the cost of the ability to give specific, intuitive interpretation of the yield curve movements. Most important, internal basis models face many empirical difficulties similar to those in the estimation of external basis models, in particular, overfitting and small-sample problems.

The remainder of this article is organized as follows. The next section reviews the standard affine-Gaussian setup of macro-finance models, derives the affine bond pricing formula in a way that emphasizes the replicating portfolio intuition, and introduces a distinction between internal basis models and external basis models. A critical examination of the assumptions in both "low-dimensional" and "high-dimensional" external basis models follows this review. Next, I then discuss the challenge of accommodating non-linear/non-Gaussian effects, such as structural breaks and time-varying uncertainties, and potential problems with empirical techniques commonly used in the estimation and evaluation of macro-finance models. I then return to *why* surveys perform better than models in inflation forecasting (as documented by ABW).

[4] For example, the root mean square errors (RMSEs) for 1-year consumer price index (CPI) inflation forecasts based on the two no-arbitrage models in ABW (what they refer to as MDL1 and MDL2) are larger than those of autoregressive (AR)(1) and autoregressive moving average (ARMA)(1,1) models by more than 30 percent for the post-1995 window. Furthermore, all but 1 of ABW's 11 regression models that involve term structure variables (what they refer to as TS1-TS11) produce smaller RMSEs in forecasting 1-year CPI (PUNEW) inflation than the no-arbitrage models in the post-1995 sample.

THE BASIC MODEL

Affine-Gaussian Framework

Most macro-finance models in the literature are based on the "affine-Gaussian" model, denoted as

(1)
$$m_{t+1} \equiv \log M_t = -r(x_t) - \lambda(x_t)' \varepsilon_{t+1} - \frac{1}{2}\lambda(x_t)' \lambda(x_t)$$
$$x_{t+1} = \Phi x_t + (I - \Phi)\mu + \Sigma\varepsilon_{t+1},$$
$$r(x_t) = \rho_o + \rho' x_t$$
$$\lambda(x_t) = \lambda_a + \Lambda_b x_t,$$

where M_t is the pricing kernel, x_t is an n-dimensional vector of state variables, r_t is the nominal short rate (i.e., one-period yield), and λ_t is the market price of risk of the n-dimensional shocks ε_{t+1} (Φ, Σ, and Λ_b are $n \times n$ constant matrices, ρ and λ_a are constant n-dimensional vectors, and ρ_o is a constant). A well-known result in finance theory states that no-arbitrage implies the existence of a pricing kernel (stochastic discount factor) of the form (1).[5]

There is freedom in choosing the specific functional form of r_t and λ_t and the dynamics of x_t. Use of the affine forms for r_t and λ_t and the Gaussian specification (VAR(1) specification) of x_t constitutes the affine-Gaussian model. This form has certain limitations (discussed later), but it is still quite general and capable of encompassing many of the known models in finance and macroeconomics.

Using the recursion relation for the price of a τ-period zero-coupon bond at time t,

(2)
$$P_{\tau,t} = E_t\left(P_{\tau-1,t+1}M_{t+1}\right),$$

it is straightforward to show that bond prices in this model are given by

(3)
$$P_{\tau,t} = \exp\left(A_\tau + B_\tau' x_t\right),$$

where A_τ and B_τ are the solution of the difference equations,

(4)
$$0 = \rho_o + A_\tau - A_{\tau-1} - \frac{1}{2}B_{\tau-1}'\Sigma\Sigma' B_{\tau-1}$$
$$-B_{\tau-1}'\left((I-\Phi)\mu - \Sigma\lambda_a\right) \text{ and}$$
$$0_{n\times 1} = \rho + B_\tau - \left(\Phi - \Sigma\Lambda_b\right)' B_{\tau-1},$$

with boundary condition $A_{\tau=0} = 0$, $B_{\tau=0} = 0_{n\times 1}$ (see, for example, Ang and Piazzesi, 2003 [AP]). The bond yield $y_{\tau,t}(=-\log(P_{\tau,t})/\tau)$ is given by

(5)
$$y_{\tau,t} = -\frac{1}{\tau}A_\tau - \frac{1}{\tau}B_\tau' x_t,$$

that is, it takes an affine form.

The original "finance term structure models" such as those by Dai and Singleton (2000) and Duffee (2002) were written for nominal bond yields only. For example, the model defined by equation (1) could be estimated with just nominal yields data, with suitable (normalization) restrictions on the parameters Φ, μ, ρ,... to ensure that the model be econometrically identified. The state variables in this case are "latent factors" without an explicit economic meaning.

In a seminal paper, AP proposed combining this setup with a description of the macroeconomy. Their basic insight is that the well-known Taylor-rule specification of the short rate also has an affine form:

(6)
$$r_t = \rho_\pi \pi_t^Y + \rho_g\, gap_t + const,$$

where π_t^Y is the annual inflation and gap_t is the gross domestic product (GDP) gap (log GDP minus log potential GDP).[6] Therefore, using variables such as inflation and the GDP gap as part of the state vector in equation (1), that is,

(7)
$$x_t = \left[\pi_t^Y, gap_t, \ldots\right]',$$

provides a system in which bond yields are linked to key macroeconomic variables. Some macroeconomic variables might not be well described by simple VAR(1) dynamics, but this is, in principle, not a problem, as a higher-order VAR process (VAR(q) model) can be written as a VAR(1) process

[5] Duffie (2001) discusses this in the continuous-time formalism; see also Cochrane (2001).

[6] To be precise, AP use GDP growth, instead of the GDP gap, in their formulation.

with an expanded state vector that includes lags of these variables (e.g., $[\pi_t^Y, \pi_{t-1}^Y, \ldots, gap_t, gap_{t-1}, \ldots]'$).

Various macro-finance models differ by the choice of the restrictions imposed on the matrices (like Φ, ρ,..., etc.). For example, AP adopt an atheoretical (statistical) approach, reminiscent of Sims's (1980) original VAR proposal; Hördahl, Tristani, and Vestin (2006; HTV) impose more structure, based on a New Keynesian model as in Clarida, Galí, and Gertler (2000), though still remain in the reduced-form framework.

These models are an innovation from the earlier approach of handling long-term bond yields in macroeconomic models. In fact, most macroeconomic models have not dealt with long-term bond yields at all, despite their importance for savings and investment decisions in the economy. Precursors to the macro-finance models, such as the Federal Reserve's FRB/US model, do contain the 5-year and 10-year nominal yields, which are specified as the expectations hypothesis–implied yield plus a term premium (the 5-year term premium and the 10-year term premium are modeled separately),[7] but the framework (equation (1)) allows not just a few selected long-term yields but information from the *entire* yield curve to be integrated with a description of the macroeconomy.

No-Arbitrage and Replicating Portfolios

Although the derivation of the affine bond pricing equation (3) using the recursion relation involving the pricing kernel is simple and elegant, it is useful to re-derive it using the hedging (spanning) argument[8] to get a better sense of the role that the no-arbitrage principle plays in macro-finance models. Suppose there are n-dimensional shocks underlying the term structure movements, denoted by a standard normal random vector ε_t. The change in the value of a bond with maturity τ can be expressed generally as

$$\frac{\delta P_{\tau,t+1}}{P_{\tau,t}} = \mu_{\tau,t} + \gamma'_{\tau,t}\varepsilon_{t+1}, \tag{8}$$

where I have used the notation $\delta P_{\tau,t+1}$ for $P_{\tau-1,t+1} - P_{\tau,t}$ (the change in the value of a bond which was of time-to-maturity τ at time t) to avoid confusion with simple time-differencing; $\Delta P_{\tau,t+1} = P_{\tau,t+1} - P_{\tau,t}$; $\mu_{\tau,t}$ is the one-period expected return on a bond that has time-to-maturity τ at time t (i.e., $\mu_{\tau,t} = E_t(\delta P_{\tau,t+1}/P_{\tau,t})$); and the n-dimensional vector $\gamma_{\tau,t}$ is the loading on the shocks that determine the unexpected return.

Consider a portfolio formed by taking positions in $n+1$ bonds with maturities $\tau_1, \tau_2, \ldots, \tau_{n+1}$, with portfolio weights $w_{1t}, \ldots, w_{n+1,t}$. Denoting the value of this portfolio, the return on the portfolio V is given by

$$\begin{aligned} \frac{\delta V}{V} &= w_1 \frac{\delta P_{\tau_1}}{P_{\tau_1}} + \ldots + w_{n+1}\frac{\delta P_{\tau_{n+1}}}{P_{\tau_{n+1}}} \\ &= \sum_{i=1}^{n+1} w_i \mu_{\tau_i} + \left(\sum_{i=1}^{n+1} w_i \gamma_{\tau_i}\right)' \varepsilon, \end{aligned} \tag{9}$$

where the time index t has been suppressed for notational simplicity. If the portfolio is locally risk-free

$$\left(\sum_{i=1}^{n+1} w_i \gamma_{\tau_i} = 0\right),$$

then by no-arbitrage it should yield a risk-free rate (one-period yield); that is,

$$\sum_{i=1}^{n+1} w_{it}\,\mu_{\tau_i,t} = r_t,$$

which is equivalent to

$$\sum_{i=1}^{n+1} w_{it}\left(\mu_{\tau_i,t} - r_t\right) = 0$$

since

$$\left(\sum_{i=1}^{n+1} w_{it} = 1\right).$$

Summarizing, we have

$$\sum_{i=1}^{n+1} w_{it}\left(\mu_{\tau_i,t} - r_t\right) = 0,\ \sum_{i=1}^{n+1} w_{it}\gamma_{\tau_i,t} = 0_{n\times 1}. \tag{10}$$

---

[7] See, for example, Brayton et al. (1997).

[8] The derivation here can be viewed as a discrete-time analogue of Cox, Ingersoll, and Ross's (1981) continuous-time derivation.

Recall that the latter of these equations is n-dimensional, since $\gamma_{\tau_i,t}$'s are n-dimensional. Equation (10) can be put in the matrix form,

$$(11)\quad \begin{bmatrix} \mu_{\tau_1,t}-r_t & \cdots & \mu_{\tau_{n+1},t}-r_t \\ \gamma_{\tau_1,t} & \cdots & \gamma_{\tau_{n+1},t} \end{bmatrix} w_t = 0_{(n+1)\times 1},$$

where $w_t = [w_{1t},\ldots,w_{n+1,t}]'$. In order for this matrix equation to have a nontrivial (i.e., nonzero) solution w_t for an arbitrary choice of τ_i's, the expected excess return $\mu_{\tau,t}-r_t$ has to be a linear combination of $\gamma_{\tau,t}$; that is,

$$(12)\quad \mu_{\tau,t}-r_t = \gamma'_{\tau,t}\lambda_t,$$

where the n-dimensional vector λ_t ("market price of risk") expresses the linear dependence between $\mu_{\tau,t}-r_t$ and $\gamma_{\tau,t}$.

It is often more convenient to deal with log prices and log returns on bonds, $\delta \log P_{\tau,t+1}$ $(=\log P_{\tau-1,t+1}-\log P_{\tau,t})$. From the discrete-time version of Ito's lemma,[9] one has

$$(13)\quad \delta \log P_{\tau,t+1} = \tilde{\mu}_{\tau,t} + \gamma'_{\tau,t}\varepsilon_{t+1},$$

where

$$(14)\quad \tilde{\mu}_{\tau,t} = E_t\left(\frac{\delta P}{P}\right) - \frac{1}{2}var_t\left(\frac{\delta P}{P}\right) = \mu_{\tau,t} - \frac{1}{2}\gamma'_{\tau,t}\gamma_{\tau,t}.$$

Thus, equation (12) can be also written

$$(15)\quad \tilde{\mu}_{\tau,t} - r_t + \frac{1}{2}\gamma'_{\tau,t}\gamma_{\tau,t} = \gamma'_{\tau,t}\lambda_t.$$

Note that the derivation thus far has been quite general. If the short rate and market price of risk are affine in the state variables and if the state variables follow a VAR(1) process (i.e., equation (1)), one obtains a particularly simple result. Positing that the bond prices take the form $\log P_{\tau,t} = A_\tau + B'_\tau x_t$, one has (from equation (13))

$$(16)\quad \tilde{\mu}_{\tau,t} = A_{\tau-1} - A_\tau + B'_{\tau-1}(I-\Phi)\mu + (B'_{\tau-1}\Phi - B'_\tau)x_t$$

$$(17)\quad \gamma'_{\tau,t} = B'_{\tau-1}\Sigma.$$

Substituting these (and the expressions for r_t and λ_t) into equation (15) gives the same difference equation for bond prices as in equation (4)—hence, the same bond prices, as promised earlier.

Internal Basis Models versus External Basis Models

The key formula in the above derivation of the bond pricing equation is equation (12), or equivalently, equation (15). It states that the expected return on a bond of arbitrary maturity in excess of the short rate depends on the product of the bond-independent market price of risk, λ_t, and the bond's sensitivity to risk, $\gamma_{\tau,t}$. The basic intuition underlying equation (12) is that *the yield curve is "smooth," so the risks to a bond can be hedged well by a portfolio of (a relatively small number of) other bonds*. This is well known from the factor analysis of Litterman and Scheinkman (1991) and other studies. One can also see this from the regression of the quarterly change in the 5-year yield on the changes in 6-month, 2-year, and 10-year yields, which gives very high R-squareds (e.g., 99 percent).

Note that equation (12) itself is silent about the structure of the λ_t vector, except for the condition that it does not depend on bond-specific information (like maturity). In fact, the early generation of affine-Gaussian models assumed a constant market price of risk vector λ, which in effect implied a version of the expectations hypothesis. Later studies recognized that λ_t can depend on the state of economy; thus, a variable influencing the market price of risk would also influence bond prices.[10] However, this creates, in a sense, too large a set of possibilities: Any variable could, in principle, enter the expression for the market price of risk and, in turn, the expression for bond yields.

Latent-factor models of the term structure, such as the affine-Gaussian model of Duffee (2002) (the $EA_0(n)$ model in Duffee's terminology), partly get around this problem by implicitly defining the model in statistical terms. A "maximally flexible" n-dimensional affine-Gaussian model (1) can be viewed as an answer to the question, "what is the most general n-dimensional representation of the yield dynamics in which yields

---

[9] See, for example, Campbell, Chan, and Viceira (2003).

[10] Thus, a shock that changes λ_t, say ξ_t, should also be included in the vector of shocks ε_t that moves bond prices.

are Gaussian, linear in some basis, and consistent with no arbitrage?" As the yield curve seems to be well described by a small number of risk sources, it stands to reason that there exists a suitable representation for a relatively small n. Thus, the no-arbitrage principle in this setting can help describe the rich variation of the yield curve in a tractable and relatively parsimonious way, while allowing for a general pricing of risk (as opposed to the expectations hypothesis).

Duffee's (2002) affine-Gaussian model describes only the nominal yield curve, but it is straightforward to write down a "joint model" of nominal yields and inflation in the same spirit by combining equation (1) with the following specification of the inflation process:

$$\begin{aligned} \pi_{t+1} &= \chi(x_t) + \tilde{\sigma}'\tilde{\varepsilon}_{t+1} \\ \chi(x_t) &= \psi_0 + \psi' x_t, \end{aligned} \tag{18}$$

where the one-period inflation π_{t+1}(=log(Q_{t+1}/Q_t), Q_t being the price level) consists of the one-period *expected* inflation $\chi(x_t)$ and unforecastable inflation $\tilde{\sigma}'\tilde{\varepsilon}_{t+1}$. As in the case of the nominal short rate r_t, the one-period inflation expectation is specified as an affine function of the state vector x_t. The disturbance vector $\tilde{\varepsilon}_t$ includes the vector of shocks that move interest rates (ε_t in equation (1)) and a shock (say ε_t^+) that is orthogonal to the interest rate shocks.[11] As in the nominal-yields-only model, the state vector x_t is a vector of statistical variables (latent variables), which is determined only up to normalization restrictions (on parameter matrices Φ, ρ_0, ρ, ψ_0, ψ,...) that ensure the (maximal) identification of the model. I shall refer to such a model as an "internal basis model," as the state vector is unknown before the estimation and is determined *inside* the estimation with yields, inflation, and possibly other data.[12]

Such a joint model makes only fairly weak assumptions: Writing the one-period inflation as the sum of expected inflation and unexpected inflation in equation (18) is quite general, and it makes intuitive sense to have the state vector x_t describe inflation expectations and bond yields together, as a variable that moves inflation expectation would also be expected to move nominal interest rates. At the same time, this formulation relaxes the assumptions implicit in the two traditional theories of nominal yields: It goes beyond the expectations hypothesis—as it now allows for time-varying term premia—and the Fisher hypothesis—as it now implicitly allows for a general correlation between real rates and inflation. Note that the state vector x_t in the joint model has more economic meaning than the nominal-yields-only model in the sense that it is now (implicitly) related to objects such as inflation expectations and inflation risk premia. However, the fact that the x_{it}'s are still latent factors is potentially an unattractive feature and makes it difficult to discuss bond market developments in a simple manner.

Thus, many papers in the macro-finance literature take all or part of the state vector to be specific macroeconomic variables (or variables with clear macroeconomic interpretation) so as to make the connection between the yield curve and macroeconomy more explicit. These variables form an external basis, in the sense that they are a priori fixed, partially ("mixed" models) or completely (observables-only models). Simply speaking, internal basis models try to project information in yields $y_{\tau,t}$ and "observable" macroeconomic variables f_{it}^o onto the state vector x_t consisting of unobservable variables f_{it}^u, while external basis models try to project information in yields onto "observable" macroeconomic variables f_{it}^o and latent variables (if there are any). Schematically,

$$\textit{internal basis}: \{y_{\tau,t}\}, \{f_{it}^o\} \Rightarrow x_t = \left[f_{1t}^u, f_{2t}^u, \ldots\right]'$$

$$\textit{external basis}: \{y_{\tau,t}\} \Rightarrow x_t = \left[f_{1t}^o, f_{2t}^o, \ldots, f_{1t}^u, f_{2t}^u, \ldots\right]'.$$

As one moves on to external basis models, one might be also moving away from the relative comfort of the original intuition behind no-arbitrage (the smoothness of the yield curve); hence, close scrutiny of the additional assumptions they involve is warranted.

[11] This shock (ε_t^+) is introduced to allow for shocks to inflation that are not spanned by interest rate shocks. (One can also define the ε_t vector in equation (1) to include this shock.)

[12] Perhaps the best-known example of internal basis models is factor analysis (e.g., Litterman and Scheinkman, 1991). As in the no-arbitrage internal basis models, the factors in factor analysis are determined only up to an invariant transformation; thus, normalization restrictions are needed to define them.

EXAMINING THE ASSUMPTIONS IN EXTERNAL BASIS MODELS

Unspanned Short-Run Inflation

One implication of having a macroeconomic variable like inflation as a state variable in the setup of equation (1) is that short-run inflation risk can be hedged by taking positions in nominal bonds.[13] Many practitioners, however, would be skeptical about this claim. Policymakers are well aware of large short-run variations in price indices such as the producer price index (PPI) and consumer price index (CPI) that do not require a policy response, and they are careful to "smooth through the noise" in interpreting data on inflation. Blinder (1997) puts this clearly and strongly: "[The noise issue] was my principal concern as Vice-Chairman of the Federal Reserve. I think it is a principal concern of central bankers everywhere."

Market participants are also (implicitly) cognizant of these issues. One evidence is the bond market's reaction to the announcement of total CPI (also called "headline CPI" or simply "CPI") and core CPI (which is an inflation measure obtained by stripping out the volatile food and energy prices from total CPI): Bond yields are known to react mainly to the surprise component of core CPI, not total CPI.[14] This raises the question whether it is reasonable to treat the fluctuation in *total* CPI as risks that are spanned by the yield curve factors (an implicit assumption in most external basis macro-finance models).

One can also consider the regression of the change in quarterly inflation onto the changes in 6-month, 2-year, and 10-year yields,[15] which gives an R^2 of at most 10 percent in the 1965-2006 period, in stark contrast to the aforementioned regression of the change in the 5-year yield (R^2 of 99 percent). Even when the lagged inflation terms are included, as in

$$\begin{aligned}\Delta\pi_t &= a + b_1\Delta\pi_{t-1} + b_2\Delta\pi_{t-2} + b_3\Delta\pi_{t-3} \\ &+ b_4\Delta y_{6M,t} + b_5\Delta y_{2Y,t} + b_6\Delta y_{10Y,t} + e_t,\end{aligned} \tag{19}$$

the R^2's do not exceed 40 percent[16]; the use of more than three yields does not make much difference. This exercise is similar in spirit to Collin-Dufresne and Goldstein (2002), who argue that the relatively low R^2's in the regressions of the changes in interest rate derivative prices on the changes in interest rates indicate the presence of "unspanned stochastic volatility" in interest rates.

Note lastly that, although we have focused on inflation (CPI) here, the concern about unspanned shocks in macroeconomic variables is more general; for example, variables such as quarterly GDP growth face similar problems.

Do Macro Variables Form a Suitable Basis for Representing Expectations?

Let us now address a related question: whether external basis models can properly describe inflation expectations, which, according to the Fisher hypothesis intuition, is an important determinant of the nominal term structure.

To those who engage in inflation forecasting extensively, the poor inflation forecast performance of macro-finance models like those of ABW might not be a surprise: A long line of research has explored the inflation forecasting performance of the yield curve information and generally obtained disappointing results. Stock and Watson (2003) summarize the situation thus: "With some notable exceptions, the papers in this literature generally find that there is little or no marginal information content in the nominal interest rate term structure for future inflation."

Most of the regression-based inflation forecasting models in the literature include current and

---

[13] Let the first element of the state vector x_t in equation (1) be inflation. The formalism (1) then implies that one can in general form a portfolio of bonds that replicates the inflation shock ε_{1t}.

[14] One can regress the change in bond yield (around the announcement) on the surprise components in total CPI and core CPI (computed as the announced number minus the Bloomberg consensus prediction). The coefficient on the total CPI surprise is found to be insignificant.

[15] Let y_t denote a vector of n yields. In the affine model, one has $y_t = a + Bx_t$, where x_t is the n-dimensional vector of state variables, a is an n-dimensional constant vector, and B is an $n \times n$ constant matrix. Inverting it gives $x_t = B^{-1}a + B^{-1}y_t$. Thus, if π_t is an element of x_t, this implies $\Delta\pi_t = c'\Delta y_t$, where c is a constant vector. This is in the linear regression form *without the residual error term* (and the intercept term).

[16] In the regression (19), I have tried quarterly inflation based on both the quarter-averaged CPI and the end-of-quarter (last month of the quarter) CPI. In these cases, the quarter-averaged yields and end-of-quarter yields were used, respectively.

lagged inflation as regressors to take into account the persistence of inflation. The expected inflation over the next year in these models takes the form

$$E_t\left(\pi_{t+1Y,t}\right)=a+b_0\pi_t^*+b_1\pi_{t-1}^*+\ldots+c'z_t, \tag{20}$$

where π_t^* is either the one-period inflation or annual inflation and the vector z_t denotes other regressors, which could include term structure variables.

Consider a macro-finance model (1) that has quarterly (one-period) inflation π_t as a state variable. In other words, $x_t = [\pi_t, \tilde{z}_t]'$, where $\tilde{z}_t(=[\tilde{z}_{1t}, \tilde{z}_{2t},\ldots]')$ denotes other state variables. The expected inflation over the next year is

$$\begin{aligned}&E_t\left(\pi_{t+1Y,t}\right)\\ &=[1,0,\ldots,0]\left(\left(\Phi+\Phi^2+\Phi^3+\Phi^4\right)\left(x_t-\mu\right)+\mu\right)\\ &=\tilde{a}+\tilde{b}_0\pi_t+\tilde{b}_1\tilde{z}_{1t}+\tilde{b}_2\tilde{z}_{2t}+\ldots,\end{aligned} \tag{21}$$

which is in the same form as equation (20).[17] (The case is similar with models that use annual inflation as a state variable.) As such, the difference between the macro-finance models formulated this way and the regression models is simply in the *coefficients*, not in the *basis*. There is a possibility of an "efficiency gain" with no-arbitrage models (through the imposition of useful constraints on the coefficients), but even this is not ensured if the results in ABW are any indication. More fundamentally, though, the frequently poor inflation forecast performance of regression models and macro-finance models like ABW raises questions about the efficacy of the basis itself.

Lessons from Simple Models

Some of the key conceptual issues in the representation of the yield curve and inflation expectations may be explained through a comparison of two simple models of inflation—namely, AR(1) and ARMA(1,1) models:

$$\pi_t=(1-\phi)\mu+\phi\pi_{t-1}+\varepsilon_t\,(\mathrm{AR}), \tag{22}$$

$$\pi_t=(1-\phi)\mu+\phi\pi_{t-1}+\varepsilon_t-\alpha\varepsilon_{t-1}(\mathrm{ARMA}). \tag{23}$$

The τ-period-ahead inflation expectations in both models take the form

$$E_t\left(\pi_{t+\tau}\right)=\phi^{\tau-1}\left(\chi_t-\mu\right)+\mu, \tag{24}$$

where the expected one-period inflation $\chi_t \equiv E_t(\pi_{t+1})$ for the AR(1) model is given by

$$\chi_t=\phi\pi_t+(1-\phi)\mu \tag{25}$$

and χ_t for the ARMA(1,1) model is given by

$$\chi_t=\phi\pi_t-\alpha\varepsilon_t+(1-\phi)\mu. \tag{26}$$

The estimate of ϕ in the AR(1) model, based on U.S. quarterly CPI inflation data from 1960:Q1 to 2005:Q4, is $0.785_{(0.045)}$, while the estimates of ϕ and α in the ARMA(1,1) model are $0.935_{(0.031)}$ and $0.341_{(0.081)}$, respectively, with standard errors listed in parentheses. These numbers imply fairly similar 1-quarter-ahead inflation expectations, as can be seen in Figure 1A. (There is somewhat more jaggedness in the AR(1) forecast.) The same parameter estimates, however, imply very different longer-horizon inflation expectations (Figure 1B): The 5-year-ahead (20-quarter-ahead) inflation expectation from the AR(1) model is almost constant, while the 5-year-ahead inflation expectation from the ARMA(1,1) model is more variable. (This reflects the difference between $0.785^{20-1} = 0.01$ versus $0.935^{20-1} = 0.28$ in equation (24).)

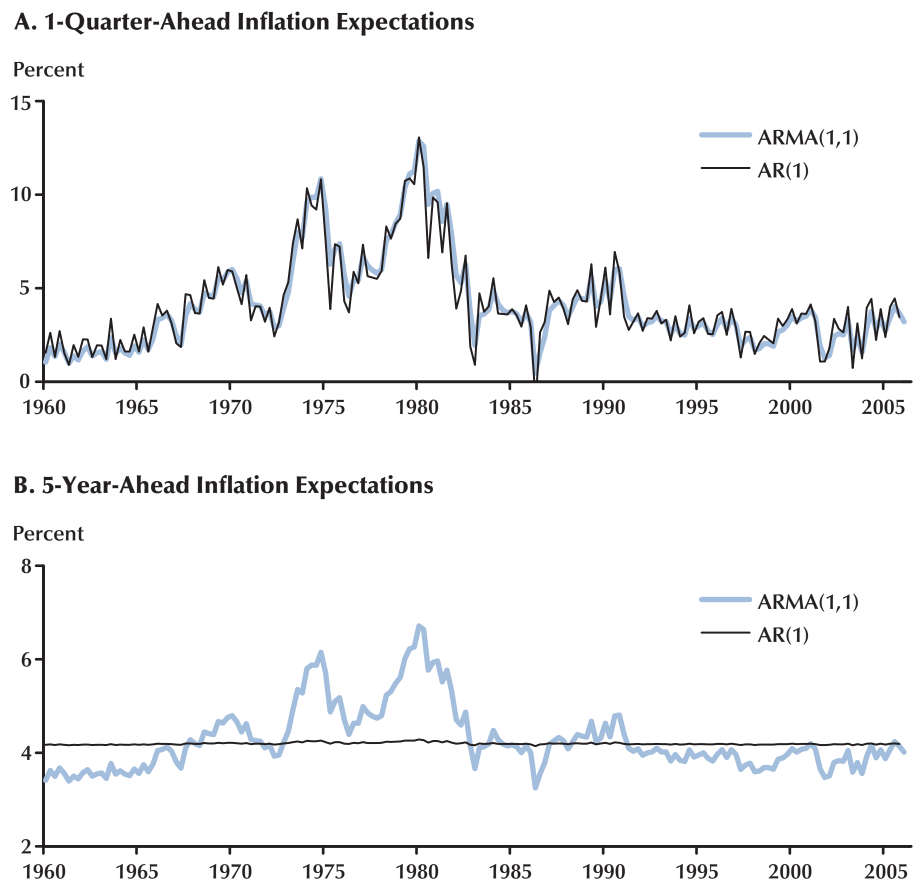
An almost constant 5-year-ahead inflation expectation from the AR(1) model in the past 40 years is implausible. The main reason for the qualitative difference between the AR and ARMA models is that the ARMA(1,1) model tries to separate the "unforecastable inflation" from the expected inflation, while the AR(1) model does not. This can be seen from the fact that the ARMA(1,1) model is a univariate representation of the following "two-component model":

$$\begin{aligned}&\pi_t=\chi_{t-1}+\eta_t\\ &\chi_t=(1-\phi)\mu+\phi\chi_{t-1}+e_t,\\ &\eta_t\sim N\left(0,\sigma_\pi^2\right),\;\; e_t\sim N\left(0,\sigma_x^2\right),\;\; corr\left(\eta_t,e_t\right)=\varrho,\end{aligned} \tag{27}$$

[17] If the $\tilde{z}_t$ vector includes latent factors, they can be "inverted" and expressed in terms of yields (because of the linear relationship between the yield and the factors), which again leads to the form of equation (20). However, latent factors in these "mixed" models are partly defined by their relation to the macro factors, and this may entail complications, as discussed later in the "Low-Dimensional External Basis Models" subsection.

Figure 1

U.S. Inflation Expectations Based on AR(1) and ARMA(1,1) Models



in which χ_t is an expected inflation process and η_t is an unforecastable inflation.[18] Though simple, this two-component model (of which the internal basis model [equation (18)], previously discussed in the "Internal Versus External Basis Models" subsection, can be viewed as an extension) is quite useful for illustrating some of the key points in this paper.[19]

[18] The MA(1) coefficient in the ARMA(1,1) model is related to the two-component model parameters as

$$\alpha = c - \sqrt{c^2 - 1}, \text{ where}$$
$$c \equiv \left[\left(1+\phi^2\right)\sigma_\pi^2 + \sigma_x^2 - 2\varrho\phi\sigma_\pi\sigma_x\right] \Big/ \left[2\left(\phi\sigma_\pi^2 - \varrho\sigma_\pi\sigma_x\right)\right].$$

See, for a derivation, Cochrane (2001, pp. 418-20).

The unforecastable inflation component η_t in equation (27) can help explain several puzzling empirical results in the literature. Among them is the negative one-lag autocorrelation of the changes in quarterly inflation $\Delta\pi_t (= \pi_t - \pi_{t-1})$, which, according to Rudd and Whelan (2006, Sec III.C), is evidence against the New Keynesian Phillips curve models (which generate positive one-lag autocorrelation). In the case of the two-component model (27), one has

[19] This model of inflation has an interesting parallel with the consumption-based asset pricing model of Bansal and Yaron (2004), who argue that writing the consumption growth Δc_t as the sum of expected component and unexpected component ($\Delta c_t = \chi_t + \eta_{t+1}$) can help resolve the equity premium puzzle.

$$\begin{aligned}
&cov\left(\Delta\pi_t,\Delta\pi_{t-1}\right)\\
(28)\quad &= cov\left(\Delta\eta_t,\Delta\eta_{t-1}\right)+cov\left(\Delta\chi_{t-1},\Delta\chi_{t-2}\right)\\
&+cov\left(\Delta\chi_{t-1},\Delta\eta_{t-1}\right).
\end{aligned}$$

The obviously negative first term dominates the second and third terms at appropriate parameter values, resulting in a negative $cov(\Delta\pi_t,\Delta\pi_{t-1})$. The unforecastable component η_t also plays the role of putting an upper bound on the predictability of inflation.

Economically, the η_t term represents very-short-run effects in total CPI inflation, including part of the food and energy prices that create the wedge between total CPI and core CPI, as well as the unforecastable components of the core CPI inflation and potential errors in the measurement of CPI.

The importance of the η_t term in the two-component model (27) has a parallel implication for no-arbitrage macro-finance models: The failure to separate out the "unspanned macro shocks" in macro-finance models may produce problems that mirror those of the AR(1) inflation model. It is worth mentioning here that Stock and Watson (2007) have also recently emphasized that separating inflation into a trend component and a serially uncorrelated shock (like η_t in equation (27)) is useful for explaining key features of U.S. inflation dynamics,[20] though they do not discuss the ramifications for macro-finance (no-arbitrage) models.

It is instructive to ask about the basic variable underlying the term structure of inflation expectations in the ARMA(1,1) model. As is clear from equation (24), the basic variable is χ_t, not realized inflation, π_t. Note that in the case of the AR(1) model, χ_t *is* π_t (up to a prefactor and an intercept), as can be seen from equation (25). This is not the case for the ARMA(1,1) model: It is straightforward to show (by solving for ε_t in equation (23) and recursively substituting into equation (26)) that χ_t in the ARMA(1,1) model depends on an infinite number of lags of π_t:

$$(29)\qquad \chi_t=(\phi-\alpha)\sum_{j=0}^{\infty}\alpha^j\left(\pi_{t-j}-\mu\right)+\mu.$$

This is in the exponential smoothing form, which has been familiar at least since the work of Muth (1960).

The expression (29) suggests that (i) the connection between realized macroeconomic variables and state variables in no-arbitrage term structure models could be complicated and (ii) the poor inflation forecasting performance of regression models and no-arbitrage models with macroeconomic variables may be a more complex issue than just a matter of having "efficient" coefficients (with conventional basis). To be sure, the state variables in nominal term structure models are not simply those that underlie the variation of inflation expectations. Factors that affect the real term structure and inflation risk premia should also be included in the nominal term structure model. However, it is not clear that these additional aspects would be any better described by macroeconomic variables.

Low-Dimensional External Basis Models

Let us now consider some specific issues that arise in external basis models with a "low-dimensional" state vector. Suppose that one has a three-factor macro-finance model in the setup of equation (1), with the state vector x_t consisting of all "observable" macroeconomic variables, say, quarterly inflation, π_t, quarterly GDP growth, g_t, and the effective federal funds rate, ff_t. The inflation expectations in this model are then linear functions of contemporaneous variables π_t, g_t, and ff_t. (To see this, simply substitute $\tilde{z}_{1t}=g_t$, $\tilde{z}_{2t}=ff_t$ in equation (21).) This type of forecast (VAR(1)) has more qualitative similarity to the AR(1) model than the ARMA(1,1) model; in particular, despite its multifactor nature, it still mixes "signal" with "noise" and can therefore be expected to inherit many of the problems with the AR(1) model.

Some of the macro-finance models in the literature, including Ang, Dong, and Piazzesi (2005; ADP) and ABW, remain in a relatively low-

[20] Stock and Watson (2007) write the U.S. inflation process for the past half century as $\pi_t=\tau_t+\eta_t$, where η_t is a serially uncorrelated disturbance term. The τ_t term (what they refer to as the trend component) can be identified as χ_{t-1} in equation (27). Stock and Watson's (2007) τ_t and η_t have time-varying volatilities, a feature which they argue is important in the inflation persistence debate.

dimensional framework but use a mix of latent factors and macroeconomic variables, but these "mixed models" may still have difficulties. Consider, for example, the ABW affine model (their MDL1 model) with quarterly inflation and two latent factors, that is, $x_t = [\pi_t, f_{1t}, f_{2t}]'$. If the latent factors f_{1t}, f_{2t} are interpreted as π_{t-1}, π_{t-2}, equation (21) takes a form similar to the smoothing form (29). However, besides the issue that two lags might not be enough, one may not have the freedom to interpret f_t's this way, as that would deprive the ability to describe other aspects of the nominal term structure (e.g., real interest rates, the time-varying risk premium, or time-varying perceived inflation target).

In the mixed models, having a macroeconomic variable like π_t as a part of the state vector may cause a distortion in the inference, as the latent factors can end up absorbing the "unspanned" variation in π_t. To illustrate this schematically, suppose that the true model of the short rate is

$$r_t = \rho\tilde{\pi}_t + f_t, \tag{30}$$

where $\tilde{\pi}_t$ is the "spanned" part of the one-period inflation π_t; that is,

$$\pi_t = \tilde{\pi}_t + e_t, \tag{31}$$

with e_t denoting the unspanned component. If one uses realized inflation π_t in place of $\tilde{\pi}_t$, then

$$r_t = \rho(\pi_t - e_t) + f_t = \rho\pi_t + (f_t - \rho e_t). \tag{32}$$

Thus, the latent factor f_t would be distorted by an amount ρe_t. Though one might be tempted to regard this simply as a redefinition of the latent factor, it would imply practical differences, such as the reduced persistence of the factor dynamics.

High-Dimensional External Basis Models

Some of the external basis macro-finance models in the literature use a fairly large number of state variables that include lagged macroeconomic variables. Many such models (including those of AP and HTV) use annual inflation π_t^Y ($= \pi_{t,t-1Y}$) as a state variable instead of one-period inflation. This may help alleviate concerns about the problem with the use of one-period inflation, since the year-on-year inflation partly "smooths out" the noise in quarterly inflation: π_t^Y can be written

$$\pi_t^Y = \sum_i w_i \pi_{t-i}, \tag{33}$$

where the weights w_i are 1/4 for i = 0,1,2,3, and 0 for $i > 3$.

Note, however, that the construction (33) automatically implies a moving average structure in π_t^Y, which suggests that the simple VAR(1) description would not be a good description of its dynamics. Thus, macro-finance models that use annual inflation as a state variable typically include additional lags, for example, AP use 12 monthly lags, in effect having a VAR(12) model. Bond yields in this case depend on a "large" set of state variables that include lagged macroeconomic variables.[21]

A problem with this type of "high-dimensional" specification is that it inherits the well-known problems of the unrestricted VAR models. In fact, AP's inflation dynamics *is* a conventional VAR. They separate the vector of relevant variables into an "observable" macro vector f_t^o and an unobservable (latent) vector f_t^u, that is, $\tilde{x}_t = [f_t^{o\prime}, f_t^{u\prime}]'$,[22] and impose the restriction that the latent factors do not affect the expectation of macroeconomic variables. Their macro vector dynamics are given by the VAR(q):

$$f_t^o = \Phi_1^o f_{t-1}^o + \Phi_2^o f_{t-2}^o + \cdots + \Phi_q^o f_{t-q}^o + c^o + \Sigma^o \varepsilon_t^o, \tag{34}$$

where q = 12. Although the parameters in the matrices $\Phi_1^o,\ldots,\Phi_q^o$ are, in principle, identified and can be estimated by ordinary least squares (OLS), this kind of unrestricted VAR is well known to suffer from overparameterization problems (which

[21] Since an invertible ARMA(1,1) model can be written as an AR model with infinite lags, the use of lagged macroeconomic variables in an external basis model may partly address the deficiency of the AR(1) model (relative to the ARMA(1,1) model) discussed in the "Low-Dimensional External Basis Models" subsection. However, identifying inflation as a state variable may still be problematic conceptually (especially in the case of one-period inflation, π_t), in view of our earlier discussion regarding the difference between χ_t and π_t.

[22] Here I have attached a tilde to x_t to clarify that this is not the full state vector. The full state vector (on which bond yields depend) in AP is larger: $x_t = [f_t^{o\prime}, f_{t-1}^{o\prime},\ldots,f_t^{u\prime}]'$.

will be discussed further in the "Empirical Issues" subsection).[23]

By having only the macroeconomic variables describe inflation dynamics, AP suppressed the possibility of the yield curve saying something about future inflation. Unfortunately, it is difficult to lift that restriction. The overparameterization problem would worsen, as the full (maximally identified) model would have an even larger number of parameters: In the specification of the state vector dynamics

$$\begin{bmatrix} f_t^o \\ f_t^u \end{bmatrix} = \begin{bmatrix} \Phi_1^o & \Phi_1^{ou} \\ \Phi_1^{uo} & \Phi_1^u \end{bmatrix} \begin{bmatrix} f_{t-1}^o \\ f_{t-1}^u \end{bmatrix} + \begin{bmatrix} \Phi_2^o & \Phi_2^{ou} \\ \Phi_2^{uo} & \Phi_2^u \end{bmatrix} \begin{bmatrix} f_{t-2}^o \\ f_{t-2}^u \end{bmatrix} + \cdots + c + \Sigma\varepsilon_t, \tag{35}$$

the matrices $\Phi_1^{ou}, \Phi_2^{ou}, \ldots$ are now nonzero and have to be estimated. Furthermore, the two-step estimation procedure that AP used is no longer applicable; hence, the estimation now involves a "one-step" optimization of a very-high-dimensional likelihood function.

For specifying external basis models that contain lags of macroeconomic variables in the state vector, it is common practice to set the coefficients of the market price of risk (Λ_b matrix in equation (1)) that load on lagged macroeconomic variables to zero (e.g., AP and HTV). Even with this restriction, the number of remaining market price risk parameters is large, and modelers often make additional ad hoc restrictions on the Λ_b matrices to reduce the number of parameters further.[24] Unfortunately, setting the Λ_b coefficients on lagged macroeconomic variables to zero may be a problematic practice. It implies that the expected excess return on a bond, $\mu_{\tau,t} - r_t$, is completely spanned by *contemporaneous* macroeconomic variables (and latent factors, if there are any). Recall, from equations (12) and (17), that

$$\mu_{\tau,t} - r_t = B_\tau' \Sigma \lambda_t. \tag{36}$$

Therefore, if λ_t does not depend on lagged macroeconomic variables, neither does the bond return premium. This means that while one has

$$y_{\tau,t} = a_\tau + b_{\tau,1}\pi_t + b_{\tau,2}\pi_{t-1} + b_{\tau,3}\pi_{t-2}\cdots, \tag{37}$$

one cannot have

$$\mu_{\tau,t} - r_t = \alpha_\tau + \beta_{\tau,1}\pi_t + \beta_{\tau,2}\pi_{t-1} + \beta_{\tau,3}\pi_{t-2} + \cdots, \tag{38}$$

that is, there is an asymmetry (in the way yields and bond risk premia depend on lagged macroeconomic variables) that was not motivated by theory. Thus, in order to cast the model in a "no-arbitrage" framework, many external basis macro-finance models may be introducing arbitrary and nontrivial assumptions about the market price of risk.[25]

AFFINE-GAUSSIAN MODELS VERSUS NON-AFFINE/ NON-GAUSSIAN MODELS

Structural Stability

One potential limitation of the general framework (1) is structural stability. To be sure, the debate about the structural stability of macroeconomic relationships is not new (see, e.g., Rudebusch, 1998, and Sims, 1998). However, it may have different ramifications for internal basis models and external basis models, and hence merits a discussion here.

Note that external basis macro-finance models have often used a framework based on the Taylor rule and VARs, but many have raised questions about the instability of these specifications.[26] One may hope that concerns about structural

[23] Models like HTV have more structure (in the form of the New Keynesian Phillips curve and IS equations), which may help alleviate overparameterization concerns, but at a possibly greater misspecification risk.

[24] For example, AP and HTV assume that Λ_b is a block-diagonal matrix (a block matrix for macro factors and a block matrix for latent factors).

[25] Duffee (2006) has also recently questioned the modeling of the term premium in the macro-finance literature, more specifically, the finding in some macro-finance papers of a strong relationship between the term premium and the macroeconomy. His point is that these studies often do not provide alternatives other than the "expectations hypothesis" (zero or constant return premium) and a term premium that depends on macro variables, leading to an exaggerated role of macro variables in term premium variation.

[26] See, for example, Clarida, Galí, and Gertler (2000) about the instability of Taylor-rule coefficients and Stock and Watson (1996) about the instability of VAR coefficients.

instability would be alleviated if latent factors are also included in external basis models. For example, a macro-finance model with a Taylor-rule–like mixed specification of the short rate (similar to ADP, 2005)

$$(39)\qquad r_t = const + \rho_\pi \pi_t^Y + \rho_g\, gap_t + f_t,$$

where f_t is a latent factor, can be written as

$$(40)\quad r_t = const + \pi_t^Y + (1-\rho_\pi)(\pi_t^Y - \pi_t^*) + \rho_g\, gap_t,$$

where π_t^* $(=-f_t/(1-\rho_\pi))$ is the time-varying inflation target. However, the factor f_t may have to play a number of other roles in the model, for instance, the interest rate smoothing term, time-varying risk premium, and so on (analogously to an earlier discussion in the "Low-Dimensional External Basis Models" subsection regarding ABW's affine model). Thus, a model written with f_t as a time-varying inflation target in mind might have some difficulty capturing the intended effect.

Furthermore, there may be instabilities other than the time-varying intercept: for instance, changes in the conditional correlation of various macroeconomic variables, changes in the persistence of the macroeconomic variables, and so on. Imagine, heuristically, a situation in which the "true" model is

$$(41)\qquad r_t = c + \rho_{\pi,t}\, \pi_t^Y + \rho_{g,t}\, gap_t,$$

that is, a Taylor-rule–like short rate with time-varying loadings on the macroeconomic variables. In this case, the two-factor affine model in which the state variables are $[\pi_t^Y, gap_t]'$ is obviously misspecified. For another example, consider a "time-varying inflation-persistence model":

$$(42)\qquad \pi_t^Y = \phi_{t-1}\pi_{t-1}^Y + c + \varepsilon_t.$$

Again, identifying π_t^Y as a state variable in an affine setting would be a misspecification.

One way to address this problem is to model these effects explicitly in non-affine/non-Gaussian models.[27] However, these models, being richer than affine-Gaussian models, may be even more susceptible to overfitting concerns and may incur a greater risk of misspecification. Alternatively, the use of an internal basis (while still remaining in the affine-Gaussian setup) may allay structural instability concerns to some extent: Internal basis models are agnostic as regards the definition of the factors; thus, a model that is obviously unstable from the point of view of an external basis may not necessarily be so from the point of view of an internal basis. For example, going back to equation (41), choosing the state variable as $x_t = [\rho_{\pi,t}\pi_t^Y, \rho_{g,t}gap_t]'$ may be more effective than having $x_t = [\pi_t^Y, gap_t]'$, although there may be an even better internal basis for the problem (depending on how the rest of the model is defined).[28]

Of course, no-arbitrage models with an internal basis should not be expected to answer all structural stability concerns. A strong structural instability may be difficult to capture even with an internal basis model, in which case it might be better to use a shorter, structurally more homogeneous sample.

Time-Varying Uncertainty

Another limitation of the affine-Gaussian models (both internal and external basis models) is that they imply homoskedastic yields, while there is copious evidence for time-varying volatility of yields.

Theoretically and intuitively, one *should* expect a relation between term structure variables and time-varying uncertainty about interest rates: To the extent that bond market term premia arise from risk, the changing amount of interest rate risk should translate to a changing term premium. It also stands to reason that at least a part of the variation in interest rate volatility is linked to the variation in the uncertainties about key macroeconomic variables. Various studies have noted that macroeconomic uncertainties (inflation, GDP, monetary policy) have declined since the Volcker disinflation, a phenomenon often dubbed the "Great Moderation."[29] One can expect this effect

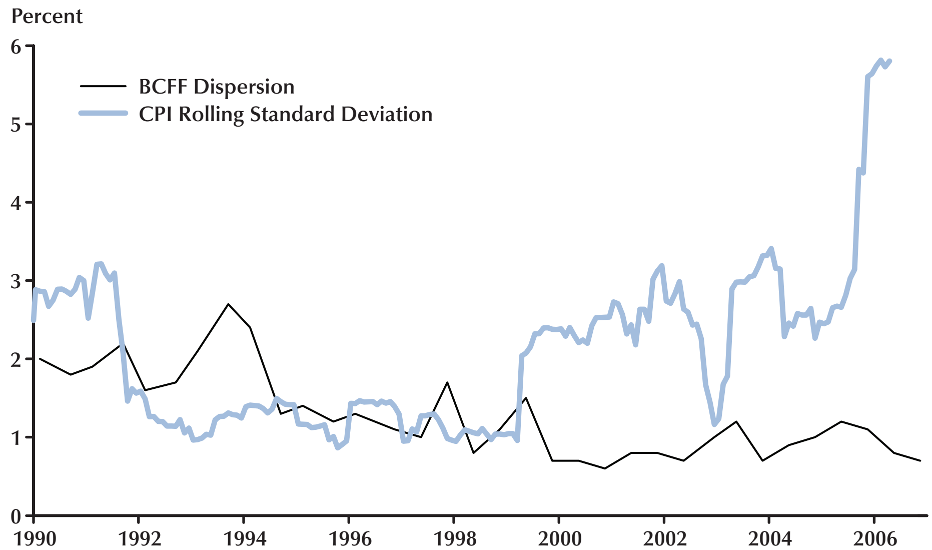
[27] For a work in this direction, see Ang, Boivin, and Dong (2007).

[28] If π_t^Y and $\rho_{\pi,t}$ are Gaussian processes, the process $\rho_{\pi,t}\pi_t^Y$ would be non-Gaussian (with time-varying volatility). However, one can still think of the affine version as an approximation of the non-Gaussian process.

[29] Bernanke (2004b) discusses this phenomenon from a policy-maker's perspective.

Figure 2

The Dispersion of Long-Horizon U.S. Inflation Forecasts in the BCFF and the 1-Year Rolling Standard Deviation of Monthly CPI Inflation



to be accompanied by a corresponding reduction in term premia in the bond market. Kim and Orphanides (2007) indeed report positive relationships between the term premium in the 10-year-forward rate and proxies for uncertainties about monetary policy and inflation based on the dispersion of long-horizon survey forecasts.[30]

However, much work remains to be done to properly address the relationship between term premia and macroeconomic uncertainties—in particular, inflation uncertainty. The key difficulty is measuring the relevant inflation uncertainty. For instance, one can debate whether the survey dispersion measure used in Kim and Orphanides (2007) is a reliable proxy for uncertainty. Inflation uncertainty measures based on a GARCH-type model also would be problematic, as they posit too tight a relationship between long-term and near-term uncertainty.[31] As can be seen in Figure 2, 1-year rolling standard deviation of monthly (total) CPI inflation (a proxy for near-term inflation uncertainty) has been elevated from around 1999 on, but this does not seem to have translated to an increase in the perception of longer-term uncertainty, proxied by the dispersion of surveyed forecasts of long-horizon inflation. Even granting the imperfection of the long-horizon inflation uncertainty measure, this contrast is noteworthy.[32]

The complexity of inflation dynamics can thus create considerable challenge for attempts to go beyond homoskedastic models: It may be that a nonlinear model with time-varying inflation uncertainty can lead to *poorer* results if the model's inflation uncertainty is misspecified, as when a model that does not make a qualitative distinction between short- and long-run inflation uncertainties tries to link the rise in the volatility of short-run inflation of the recent several years

[30] See also Backus and Wright (2007).

[31] Consider a GARCH specification of one-period inflation, $\pi_{t+1} = f(\pi_t, \pi_{t-1}, \ldots) + \varepsilon_{t+1}$, $\varepsilon_{t+1} \sim N(0, \sigma_t^2)$, $\sigma_t^2 = \alpha + \beta\sigma_{t-1}^2 + \gamma\varepsilon_t^2$. It is straightforward to show that the uncertainty about multiperiod inflation $\pi_{t+\tau,t} = (\pi_{t+1} + \pi_{t+2} + \cdots \pi_{t+\tau})/\tau$, has similar qualitative time variation as σ_t (short-run inflation uncertainty).

[32] Interestingly, earlier literature including Ball and Cecchetti (1990) and Evans (1991) has also emphasized in another context the need to distinguish between the short-run and long-run inflation uncertainties.

(as seen in Figure 2) to bond market term premia.[33]

EMPIRICAL ISSUES

Overfitting Problems

Flexibly specified no-arbitrage models tend to entail much estimation difficulty. Part of the problem is one that is familiar from unrestricted VAR models explored in the 1980s macroeconomics literature. It is well known that unrestricted VARs often lead to poor results,[34] as these models get easily overparameterized, and as the "atheoretical" (statistical) nature of these models means that there is little structure in the model to prevent estimations from generating unreasonable outcomes.

The no-arbitrage macro-finance models have two features that exacerbate the difficulty. First, unlike the unrestricted VARs, the estimation of macro-finance models typically requires nonlinear optimization (instead of OLS), because of the nonlinear relationship between bond yields and parameters such as the market price of risk and because of the latent factors (recall the discussion regarding equation (35)).

Second, the overparameterization problem can be exacerbated by a large number of additional parameters unique to macro-finance models, including those describing the dynamics of latent factors and their interaction with observable factors (e.g., Φ_j^{ou}, Φ_j^{u}, Φ_j^{uo} in equation (35)) and those describing the market price of risk (λ_a and Λ_b in equation (1)). Note that the no-arbitrage principle guarantees the existence of a pricing kernel, such as equation (1), but the principle by itself does not constrain the parameters of the market price of risk (Λ_b matrix). Suppose, as in AP, that one has in the state vector p observable variables, its $q-1$ lags, and m unobservable (latent) variables, in other words,

$$x_t = \Big[f_{1t}^o, .., f_{pt}^o, f_{1t}^u, .., f_{mt}^u, f_{1,t-1}^o, .., f_{p,t-1}^o, \cdots, f_{1,t-q+1}^o, .., f_{p,t-q+1}^o \Big]'. \tag{43}$$

In that case, the Λ_b matrix in equation (1) can have as many as $(p+m)\cdot(p\cdot q+m)$ parameters.[35] For $p=2$, $q=12$, $m=3$ (as in AP), there are 135 parameters for Λ_b to be determined; even if one chooses a smaller q, the number of parameters is still quite large.

Recall that the key innovation of the macro-finance models like AP, as compared to the traditional macro models, is that they link not only the short rate $r_t(=y_{1,t})$ but also the rest of the term structure ($\{y_{\tau,t}\}_{\tau>1}$) to the macroeconomy by casting the problem in the no-arbitrage framework (1). However, the fact that yields of various maturities tend to be highly correlated (giving rise to the finding in factor analysis and principal components analysis that there is a single dominant factor) means that the pure additional information in longer-term yields (beyond what is in the short rate) may be modest in amount and perhaps too delicate to capture with a specification of the market price of risk that is liable to be overfitted; the relation that one might see between yields and macroeconomic variables in macro-finance models may be more of a statement of the Taylor rule (macro description of the short rate) than no-arbitrage models.

The overparameterization problem may be particularly severe with external basis models that contain lags of macroeconomic variables. However, internal basis models (which tend to be implemented with a comparatively smaller number of factors, e.g., three factors) may also face serious overfitting concerns, because of the especially flexible nature (the definitional freedom) of the latent-factor models. In particular, latent factor models can do a good job of fitting the *data that they are asked to fit*, even if the model or the data are poor. For example, because yield-fitting errors are minimized as a part of the

[33] There has been much discussion about the low level of term premia in recent years (see, e.g., Backus and Wright, 2007). Trying to explain the *low* term premium and *high* uncertainty would be a daunting prospect.

[34] See, e.g., Todd (1984) and Stock and Watson (2001).

[35] The lagged macro variables do not have market price of risk, but the market price of risk of contemporaneous variables and latent variables can still depend on them, as discussed in the "High-Dimensional External Basis Models" subsection.

estimation process, internal basis models with three or four factors can fit the cross section of the yield curve quite well (with much smaller fitting errors than external basis models), but that by itself might not be a sufficient reason to recommend internal basis models.

Small-Sample Problems

The implementation of macro-finance models is also complicated by small-sample problems that arise from the highly persistent nature of the data. Both interest rates and inflation are known to be persistent; unit root tests often fail to reject a nonstationarity (unit root) null for them.

In light of this, many practitioners often use nonstationary models to forecast inflation. For example, many of the inflation forecasting models used by the Federal Reserve staff impose the unit root condition.[36] By the Fisher-hypothesis intuition, unit root inflation dynamics implies unit root interest rate dynamics.

By contrast, most of the estimated macro-finance models (or nominal term structure models) in the literature assume stationarity. Stationarity has an intuitive appeal: We do not expect interest rates and inflation to have infinite unconditional moments. Thus, we may posit that the "true model" of yields is a stationary one, perhaps with many factors to describe the complex dynamics of yields and expectations; schematically,

$$y_{\tau,t} = f_\tau\left(x_{1t}, x_{2t}, x_{3t}, \ldots., x_{Nt}\right). \tag{44}$$

In practice, however, one is forced to deal with relatively low-dimensional models, because either the limited amount of data makes it impossible to pin down the parameters of such a model or one does not have enough knowledge to construct a very detailed model. In this case, it is not clear whether the "best" *low-dimensional approximation*

$$y_{\tau,t} \approx \tilde{f}_\tau\left(\tilde{x}_{1t}, \ldots, \tilde{x}_{nt}\right), \left(n << N\right) \tag{45}$$

of the model (44) should be a stationary or nonstationary (unit root) model.

The distinction between stationary and nonstationary models could be semantic in the sense that a stationary model that is close to the unit root boundary is almost indistinguishable from unit root models, but whether to assume stationarity or not can make a big difference *operationally*, as conventional estimations have the tendency to *bias down* the persistence of stationary time series, the bias becoming stronger as the sample gets smaller. This makes the expectations appear to converge to a long-run level faster than they actually do; thus, longer-horizon expectations of inflation and interest rates in (estimated) stationary models are often artificially stable, varying little from the sample mean of these variables.

Another manifestation of the small-sample problem (besides bias) is *imprecision*: Highly persistent interest rates effectively make the size of the sample "small"; no matter how frequently the data are sampled, some of the key aspects of the term structure model (those pertaining to expectations in the physical measure, as opposed to the risk-neutral measure) are difficult to estimate.[37]

Problems with the Classical Approach

Most implementations of macro-finance models have relied on classical methods such as the maximum likelihood estimation (MLE) and generalized method of moments (GMM), but these methods may be less effective in this context than is often presumed. At the heart of the matter is the fact that reduced-form macro-finance models are obviously an *approximate representation* of data, and hence not very compatible with the classical premise of having the "true model." Though it goes without saying that all models in finance are approximate, this point is particularly relevant here in view of the atheoretical (statistical) nature of the model and the large number of parameters; the MLE or GMM criterion function of these models might thus contain multiple maxima, which capture different aspects of data with differing degree of emphasis. The small-sample problems discussed above add to the difficulty, as they make asymptotic statistics a poor guide to finite sample properties.

---

[36] Federal Reserve staff make inflation forecast judgmentally, but they do look at a variety of models to inform their judgments. The staff's forecasting procedure is discussed, for example, in Kohn (2005).

[37] The bias and imprecision problems in term structure model estimations are discussed in detail in Kim and Orphanides (2005).

Note also that many of the classical estimation approaches implicitly minimize fitting errors for the one-period-ahead conditional moments. For example, the MLE estimation can be viewed as minimizing the one-period prediction errors or the errors in the fit of the "likelihood score moments,"

$$\left(\tfrac{\partial}{\partial\theta}\log f\left(y_t \mid y_{t-1},\theta\right)\right),$$

in a GMM framework. While in theory this could yield an asymptotically correct estimate of the true model (if the true model exists), the inherently approximate nature of the model means that fitting the one-period moments as closely as possible might come at the expense of other aspects of the model. Cochrane and Piazzesi (2006) in effect make this point when they note that conventionally estimated affine models may have difficulty producing the kind of term premia that they find based on regressing multiperiod (1-year) excess returns on a set of forward rates. Note also that a GMM approach that matches unconditional moments, such as

$$E\left(y_{\tau,t}\right)=\tfrac{1}{T}\sum_{t=1}^{T} y_{\tau,t},$$
$$E\left(y_{\tau,t}y_{\tau',t-j}\right)=\tfrac{1}{T}\sum_{t=1}^{T} y_{\tau,t}y_{\tau',t-j},$$

might not be effective, due to the closeness of the interest rate process (and inflation process) to the unit root behavior.

How Can We Evaluate Models?

The above discussion suggests that looking at the fit of the moments that are often used in the classical estimation might not necessarily be a good criterion for model evaluation. Some papers *do* look directly at practical implications of the model, such as the multiperiod forecasts of inflation and interest rates. Indeed, in view of the fact that the second-moment aspects of affine-Gaussian models are trivial, much attention has focused on these conditional first moments (the forecasting performance) as a part of diagnostic criteria, as in Ang and Piazzesi (2003), HTV, and Moench (2008).

However, it is unclear to what extent summary measures of forecasting performance examined in these papers can help with model evaluation/selection. To be sure, looking at the forecasting performance can be useful for detecting problematic models. In Duffee (2002), for example, interest rate forecast RMSEs that are substantially larger than the random-walk benchmark were used to highlight problems with certain stochastic-volatility no-arbitrage models (e.g., the $EA_2(3)$ specification). Similarly, the inflation forecast RMSEs based on ABW's no-arbitrage models that are substantially larger than the univariate inflation model benchmark may signal problems with the no-arbitrage models that they have used.

Nonetheless, the RMSE measures for in-sample or out-of-sample forecasts are often ineffective in discriminating between models. For instance, ABW obtain very similar RMSEs for the 1-year out-of-sample inflation forecasts from the AR(1) and the ARMA(1,1) models, although the AR(1) model implies qualitatively quite different inflation expectations than the ARMA(1,1) model, as discussed in the "Lessons from Simple Models" subsection.

Furthermore, because a large part of the inflation and interest rate variations are unforecastable, the RMSEs themselves may have substantial uncertainty (sampling variability).[38] Thus, it may happen that the "true model" generates an RMSE that is no smaller than some other models. In this sense, it may be actually misleading to focus on the RMSE as a criterion for selecting the model that best describes reality. With in-sample forecasts, this problem is exacerbated by the possibility that RMSEs are artificially pushed down because of the use of "future information" in generating the forecast, thus making interest rates and inflation look more forecastable than they actually are.

Often there are cases in which classical criteria cannot easily tell if a model's output is unreasonable, while practitioners can do so using "judgmental information." For instance, many macro-finance models estimated with data going back to 1970s generate current (circa 2006) long-horizon inflation expectations that exceed 4 percent.

[38] Clark and McCracken (2006) emphasize that out-of-sample inflation forecast RMSEs may have weak power.

(Recall also the AR and ARMA model outputs in Figure 1B.) Though long-horizon expectations are difficult to evaluate on purely econometric grounds, as there are not many nonoverlapping observations, most policymakers and market participants would immediately say that a 4 percent long-horizon CPI inflation expectation is too high; hence, models with such an output may fail the test of relevance before any statistical tests. Note also that even if two models generated similar forecast RMSEs, practitioners could have a very different assessment of them, depending on the details of the forecast errors from the models (such as the direction of the errors).[39]

These discussions highlight the role of the larger information set of practitioners (as compared with academic researchers). Unfortunately, much of this extra information is difficult to cast in the formal language of statistical tests, and the proper evaluation of models remains a challenge for macro-finance modeling.

Would a Bayesian Approach Help?

The use of Bayesian techniques to address problems with conventional (classical) estimation has a long history, but a particularly relevant early example is the Bayesian approach to VAR forecasting. As discussed in the "High-Dimensional External Basis Models" and "Overfitting Problems" subsections, unrestricted VARs share some of the key problems encountered in flexibly specified macro-finance models, in particular, the statistical (atheoretical) nature of the specification and the tendency for overparameterization. Litterman (1986) and others have documented that a Bayesian implementation with an informative prior ("random-walk prior") can generate better results than the classical implementation. This encourages us to take up a Bayesian strategy to address the empirical difficulties with macro-finance models.

In the macro-finance context, ADP have in fact already proposed a Bayesian approach, but it is not clear that the particular priors that they have used would help overcome the problems with classical estimation discussed above. ADP state that, except for the condition that the model be stationary, their priors are *uninformative*. However, to the extent that the main problem with the classical estimation of macro-finance models is that the data by themselves are not fully informative about the model (especially as regards the overfitting and small-sample problems), it is difficult to see how uninformative priors would solve the problem. Recall that the superior performance of Bayesian VARs (over conventionally estimated unrestricted VARs) came from having an *informative* prior.

When ADP (2005) tried to estimate their model using a classical method (maximum likelihood estimation), they found that the estimated model explained most of the term structure movements in terms of the latent factor, and left little role for macroeconomic variables to explain yield curve movements,[40] an outcome that is unappealing from the viewpoint of making a connection between the macroeconomy and the yield curve. However, even granting the problems with classical methods, there may be a reason for this—namely, that the estimation marginalizes the macroeconomic variables to avoid the counterfactual implication that shocks to inflation (and other macroeconomic variables) have a tight relation to yield curve movements. This is a specification issue (i.e., one has to deal with "unspanned" variation in macroeconomic variables in the model.) Addressing the problem purely as an estimation issue may lead to problems elsewhere in the model.

In my view, the main challenge for a Bayesian implementation is in coming up with suitable

[39] For example, in the 1990s, inflation data often came in on the "low" side, and it is widely believed that not all of this had been predicted by market participants, that is, the "true" market forecast of inflation in this period likely contained a mild *upward bias*. (See Kohn, 1999, and Croushore, 1990, for Fed staff and private sector forecasts in the 1990s, respectively.) Though an "unbiased" multiperiod forecast is often viewed as a consequence of rational expectations, to obtain it one needs tight assumptions that are difficult to justify in reality—in particular, the assumptions that there is a relatively simple, structurally stable model of the economy and that the agents fully know this structure. More realistic rational expectations hypotheses that relax these restrictions, for example, models that allow for learning and time-varying structure, are consistent with biased expectations in "small" samples.

[40] Private conversation with Andrew Ang. ADP's paper (2005) itself does not describe the specifics of the outcome from the classical estimation of their model.

informative priors. This is particularly the case when there are latent factors in the model (external basis models with latent factors or internal basis models): Because the economic meaning of many of the individual parameters related to the latent factors is unclear, it is difficult to provide sensible priors for them. Recall that a flexibly specified latent factor model can be normalized in many different (but equivalent) ways. It would be problematic if a Bayesian prior that was stated for one normalization of the model did not hold in another normalization of the model.[41]

By stating priors about the variables that have direct economic meaning, such as inflation expectations, interest rate expectations, and expected bond returns, one can get around the problem of normalization dependence: Surely these variables must be normalization independent. Recall also that the source of the small-sample problem is the difficulty of estimating the parameters related to expectations (in the physical measure); thus, imposing priors on these variables would help alleviate the problem. A prior about the 10-year inflation expectation, for example, can be expressed as

$$(46)\quad i^{e}_{10Y,t} = a_{10Y}(\theta) + b(\theta)'_{10Y}\, x_t(\theta) \sim N\left(\mu_t, \sigma_t^2\right),$$

with θ denoting the model parameters collectively. For μ_t, one can use a survey median forecast. Setting $\sigma_t = \infty$ corresponds to having no priors on i^{e}_{10Y}. Setting σ_t at a large value, but not large enough to be irrelevant, can be viewed as a quasi-informative prior. Other Bayesian priors that are based on economic concepts and mechanisms may be also useable.[42]

A statement like equation (46) can be conveniently incorporated within a Kalman-filter setting. Running a Kalman-filter–based MLE with a survey median (or mean) forecast (of interest rates and/or inflation) as a noisy proxy, as in D'Amico et al. (2008), can be viewed as a "poor man's Bayesian" implementation, with the point estimate serving as the mode of the Bayesian posterior.

UNDERSTANDING THE SUPERIOR PERFORMANCE OF SURVEY FORECASTS

The specification and implementation problems discussed so far may help explain why macro-finance models, which use *more* information than past inflation data, could generate *poorer* results than simple univariate inflation models. But is the yield curve information useful at all for inflation forecasting? Why do survey forecasts perform better than univariate models (and other models)?

One reason ABW offer for the superior performance of survey forecasts is that survey participants have more information about the economy than econometricians. This is in line with the point made in the "How Can We Evaluate Models?" subsection that informational differences may create a wedge between a practitioner's and an academic researcher's evaluation of a model. But it is worth exploring this issue further.

One could plausibly expect that survey forecasts may have advantages at least at short horizons, in that a potentially vast amount of information that is relevant for forecasting near-term inflation may not be easily summarized into a small number of variables. Thus, it may be instructive to examine the near-term expectations in surveys and how they are linked to longer-term expectations (i.e., the *term structure* of survey inflation forecasts).

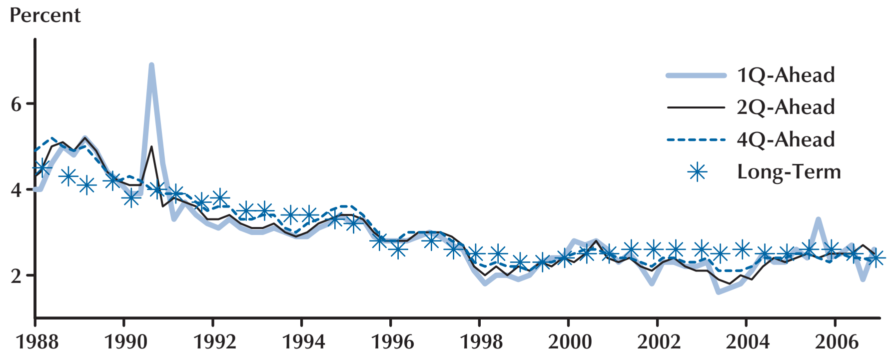
Fortunately, fairly detailed information about the near-term term structure of survey inflation expectations can be obtained, as survey forecasts such as the Survey of Professional Forecasters (SPF) and the Blue Chip Financial Forecasts (BCFF) provide CPI inflation forecasts up to the next four or more quarters. Figure 3 shows the 1-, 2-, and 4-quarter-ahead CPI inflation forecasts from the BCFF survey, based on the surveys published in January, April, July, and October (taken at the end of December, March, June, and September), from 1988 to 2006. The figure also shows the BCFF long-horizon forecast (inflation

[41] See the working paper versions of this paper (Kim, 2007) for elaboration.

[42] "Structural" priors can be also imposed in a Bayesian setting, as in the dynamic stochastic general equilibrium (DSGE) modeling literature.

Figure 3

Forecasts of U.S. Quarterly CPI Inflation from the BCFF



expected between the next 5 and 10 years), which is available twice per year, is also shown. It is notable that this long-horizon forecast, which can be viewed as a "quasi–long-run" mean of inflation, has moved about (shifted down) significantly. It is also notable how quickly the multiperiod forecasts approach the quasi–long-run value. The 4-quarter-ahead forecast and the 2-quarter-ahead forecast are already quite similar to the long-horizon forecast. Note that even in 1990:Q3, when the 1-quarter-ahead inflation expectation peaked, the expectations for longer horizons show that the survey participants expected inflation to come down quickly to the quasi–long-run level. Thus, one comes to a conclusion that "the long term is quite near."

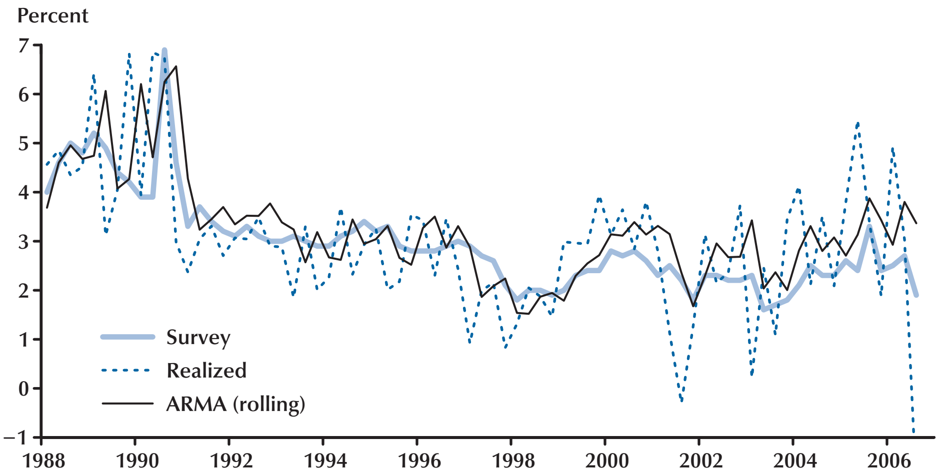
To get further insights into the survey forecasts, it is useful to compare them with ex post realized inflation and the real-time forecasts from the ARMA(1,1) model. Figure 4A shows the 1-quarter-ahead inflation forecasts based on the BCFF survey and the ARMA(1,1) model (20-year rolling sample forecast), as well as the realized quarterly inflation (π_t plotted at t–1). The vertical difference between realized inflation and the survey forecast or the ARMA forecast is the forecast error. This error is indeed smaller for the survey forecast. (The RMSEs of the 1-quarter-ahead forecast are 1.19 percent and 1.40 percent in annual percentage units for the survey forecast and the ARMA(1,1) model, respectively.) Note that the 1-quarter-ahead survey forecast is much less jagged than realized inflation or the ARMA(1,1) forecast. Granting the caveat that surveys might not necessarily be the best possible means of forecasting, this suggests that a substantial part of short-run inflation is unforecastable ex ante, lending support to a formulation like the two-component model in equation (27) in which the inflation process is separated into a trend inflation component and an unforecastable component.

Let us now examine the 1-year inflation forecast, shown in Figure 4B. The ARMA forecasts (both the rolling and the expanding samples) performed worse than the survey forecast with RMSEs of 1.04 percent for the 20-year rolling sample ARMA, 1.15 percent for the expanding sample ARMA, and 0.76 percent for the survey. The basic reason for the superior forecast of the survey is that the ARMA model–based forecasts substantially overpredicted inflation in the 1990s. It can be seen that the ARMA forecasts lie notably above the realized inflation (and survey forecast). This overprediction is due in large measure to the fact that the ARMA model in real time tended to generate "too high" values of the long-run mean level (μ in equation (23)) to which the forecasts are converging. This is illustrated in Figure 5, where the long-run mean parameter μ from the expanding sample estimation lies significantly above the long-horizon survey forecast. Because the expand-

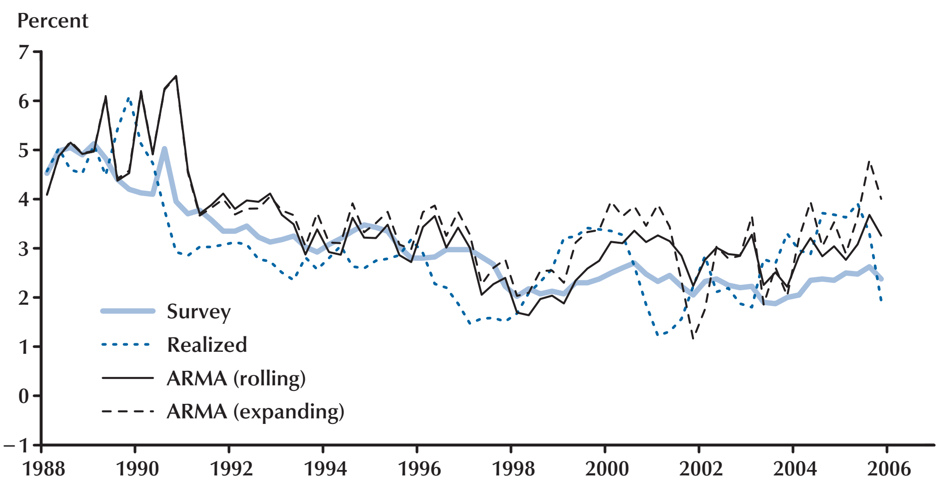
Figure 4

Comparison of Realized CPI Inflation, the BCFF Survey Forecasts, and the ARMA(1,1) Model Forecasts

A. 1-Quarter-Ahead Forecasts



B. 1-Year-Ahead Inflation Forecasts

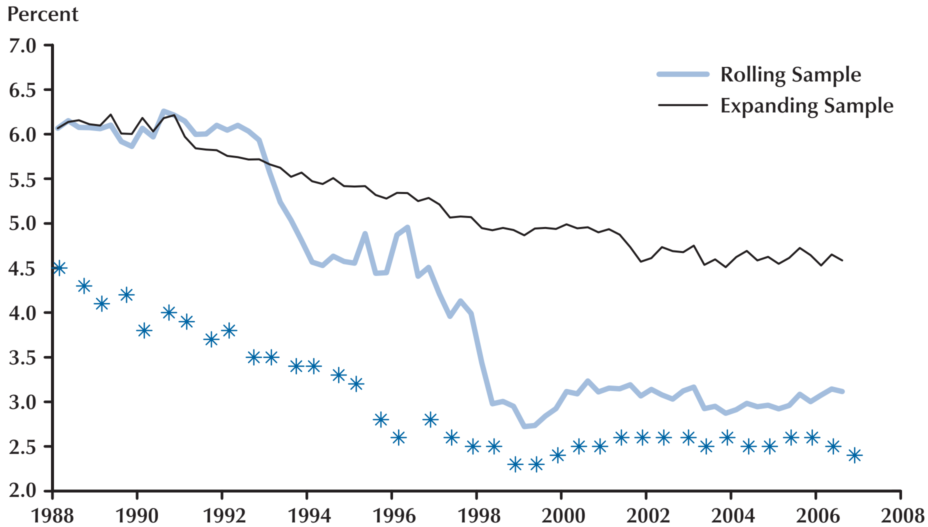


ing sample includes periods of high inflation (1970s and early 1980s), the estimated mean does not fall quickly with declining inflation in the '80s and '90s. The use of the 20-year rolling sample produces lower μ (than the expanding sample) as the estimation sample moves away from those periods, but still the adjustment in the long-run mean is not as fast as in the survey forecast.[43]

[43] ABW also note that the ability of survey forecasts to quickly adapt to major changes in the economic environment contributes to the superior performance of the surveys. While the majority of ABW's estimations were done with expanding samples, they also examine the forecast RMSEs based on rolling-sample estimation for a subset of their models. Because their rolling sample (10 years) is shorter than the 20-year rolling sample used here, ABW's rolling-sample results are even closer to those of the surveys. For example, the ratio of the AR model RMSE and the survey RMSE in the post-1995 window is 0.879/0.861, very close to 1.

Figure 5

Long-Run Means from the ARMA(1,1) Model Estimations of U.S. Quarterly Inflation Data



NOTE: The long-horizon BCFF survey inflation forecast is also shown (*).

The key point that emerges from this discussion is that surveys produce a more successful forecast of inflation in large part because they capture the *trend* component of inflation better than time-series models such as the ARMA(1,1) model. In stationary time-series models (for example, the models in Figure 4), forecasts tend to converge to a value close to the sample mean, while nonstationary models put too much weight on the recent past; thus, there is scope for judgmental information to play a role, especially if trend inflation varies significantly over time. These considerations shed light on the attention that policymakers pay to long-term inflation expectations (a better indicator of the trend inflation than realized inflation) and on the use of judgmental forecasts at central banks such as the Federal Reserve.

The importance of modeling the variation of long-term expectations deepens the challenge for macro-finance models: Besides the specification challenge, the nearly nonstationary nature of the inflation process indicated by the substantial variability of long-term survey forecasts poses considerable empirical difficulties (discussed in the previous section). These challenges notwithstanding, the discussions in this paper can be viewed as encouraging for attempts to use term structure models to extract inflation expectations: It makes intuitive sense that the yield curve contains, at least, information about trend inflation, and the indication that the near-term informational advantage of surveys seems to wear out quickly (beyond a few quarters) gives some hope that models could capture much of the variation in inflation expectations and compete with surveys.[44]

[44] Although ABW find that survey forecasts cannot be improved by combination with models that they consider, few policymakers would regard survey forecasts as the ultimate measure of inflation expectations. Consider, for example, the fact that between 1999 and 2006 the 10-year CPI inflation expectation from the SPF survey has been almost stuck at 2.5 percent. While there is a broad consensus that long-term inflation expectations were "better anchored" in the 2000s than in the earlier decades, it may be a stretch to regard that long-term inflation expectation has become so well anchored as to be practically immovable. This may be one example in which the yield curve contains useful information that is unavailable in the SPF survey.

CONCLUSION

These are some of my key points made in this paper: (i) Not all of the variation in key macroeconomic variables is related to yield curve movements. (ii) The yield curve contains useful information about the *trend* component of inflation. (iii) The no-arbitrage principle might not be sufficient to guarantee sensible outputs from macro-finance models in practice.

As I have stressed in the second section of the paper ("The Basic Model"), the spanning argument is the basis of the no-arbitrage framework; hence the presence of a short-run inflation component that is not related to yield curve movements may undermine the validity of the models that use inflation as a state variable. Such a component may also cause special difficulties when one tries to go beyond the affine-Gaussian setup to model time-varying uncertainties about macroeconomic variables explicitly. For example, as discussed in the "Time-Varying Uncertainty" subsection, monthly CPI inflation in recent years has been more volatile than in the 1990s, but there is no strong evidence that this is reflected in the yield curve (e.g., as an increased term premium); an attempt to link them may thus lead to more serious specification errors.

I have also argued in this paper that much of the "spanned" component of inflation (the part of inflation that *is* related to the yield curve) is about the *trend* component (whose importance was stressed in the discussion in the previous section of why surveys perform better). This can help resolve the puzzle that the "conventional wisdom" that the change in nominal yields often reflects changes in inflation expectations dies hard, despite the poor performance of inflation forecasting models involving term structure variables. In some sense, the latent factor models can be viewed as a way to represent markets' implicit processing (filtering) of information.

No-arbitrage models of the term structure have been viewed as a promising way to go beyond the restrictive assumptions implicit in the expectations hypothesis (about how risk is incorporated in the yield curve). However, reduced-form affine-Gaussian no-arbitrage models with flexible specification of the market price of risk can quickly become "too unrestrictive," with a profusion in the number of parameters. In other words, the no-arbitrage principle by itself may be too weak to provide enough discipline in the model. Note also that the two technical problems with estimation discussed in the "Empirical Issues" subsection (overfitting and small-sample problems) can be viewed as an extension of the specification discussion, as the main source of the problems can be viewed as insufficient information in the data or an incomplete structure in the model. For further progress, it would be desirable to come up with an effective and non-ad hoc structure on the market price of risk and other parameters of macro-finance models—or to come up with a new, intuitively appealing way to represent the term structure.

REFERENCES

Ang, Andrew; Bekaert, Geert and Wei, Min. "Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?" *Journal of Monetary Economics*, 2007, *54*, pp. 1163-212.

Ang, Andrew; Bekaert, Geert and Wei, Min. "The Term Structure of Real Rates and Inflation Expectation." *Journal of Finance*, April 2008, *63*(2) pp. 797-849.

Ang, Andrew; Boivin, Jean and Dong, Sen. "Monetary Policy Shifts and the Term Structure." Working paper, October 2007; http://neumann.hec.ca/pages/jean.boivin/mypapers/abd.pdfworking paper.

Ang, Andrew; Dong, Sen and Piazzesi, Monika. "No-Arbitrage Taylor Rules." Working paper, European Central Bank, August 2005; www.ecb.int/events/pdf/conferences/ECB-BIS_2005/ADP2004.pdf.

Ang, Andrew and Piazzesi, Monika. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics*, May 2003, *50*(4), pp. 745-87.

Backus, David and Wright, Jonathan H. "Cracking the Conundrum." Working Paper 2007-46, Federal Reserve Board Finance and Economics Discussion

Series, 2007; www.federalreserve.gov/Pubs/feds/2007/200746/200746pap.pdf.

Ball, Laurence; Cecchetti, Stephen G. and Gordon, Robert J. "Inflation Uncertainty at Short and Long Horizons." *Brookings Papers on Economic Activity*, 1990, *1990*(1), pp. 215-54.

Bansal, Ravi and Yaron, Amir. "Risks for the Long Run: A Potential Resolution of the Equity Premium Puzzle." *Journal of Finance*, August 2004, *59*(4), pp. 1481-509.

Bernanke, Ben S. "What Policymakers Can Learn from Asset Prices." Speech before the Investment Analyst Society of Chicago, April 15, 2004a; www.federalreserve.gov/BoardDocs/Speeches/2004/20040415/default.htm.

Bernanke, Ben, S. "The Great Moderation." Speech at the meetings of the Eastern Economic Association, Washington, DC; February 20, 2004b; www.federalreserve.gov/boarddocs/speeches/2004/20040220/default.htm.

Blinder, Alan S. Commentary on "Measuring Short-Run Inflation for Central Bankers." Federal Reserve Bank of St. Louis *Review*, May/June 1997, *79*(3), pp. 157-60; http://research.stlouisfed.org/publications/review/97/05/9705ab.pdf.

Brayton, Flint; Mauskopf, Eileen; Reifschneider, David; Tinsley, Peter and Williams, John. "The Role of Expectations in the FRB/US Macroeconomic Model." *Federal Reserve Bulletin*, April 1997; www.federalreserve.gov/pubs/bulletin/1997/199704lead.pdf.

Campbell, John Y.; Chan, Yeung, L. and Viceira, Luis M. "A Multivariate Model of Strategic Asset Allocation." *Journal of Financial Economics*, January 2003, *67*(1), pp. 41-80.

Campbell, John Y.; Lo, Andrew W. and MacKinlay, A. Craig. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press, 1996.

Clarida, Richard H.; Galí, Jordi and Gertler, Mark. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics*, January 2000, *115*(1), pp. 147-80.

Clark, Todd and McCracken, Michael W. "The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence." *Journal of Money, Credit, and Banking*, 2006, *38*(5), pp. 1127-48.

Cochrane, John H. *Asset Pricing*. Princeton, NJ: Princeton University Press, 2001.

Cochrane, John H. and Piazzesi, Monika. "Decomposing the Yield Curve." Presented at the September 2006 Brookings Papers on Economic Activity Conference. Working paper (March 13, 2008, revision); http://faculty.chicagobooth.edu/john.cochrane/research/Papers/interest_rate_revised.pdf.

Collin-Dufresne, P. and Goldstein, Robert S. "Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility." *Journal of Finance*, August 2002, *57*(4), pp. 1685-730.

Cox, John C.; Ingersoll, Jonathan E. and Ross, Stephen A. "A Reexamination of Traditional Hypothesis about the Term Structure of Interest Rates." *Journal of Finance*, September 1981, *36*(4), pp. 769-99.

Croushore, Dean. "Low Inflation: The Surprise of the 1990s." Federal Reserve Bank of Philadelphia *Business Review*, July/August 1998, pp. 3-12; www.philadelphiafed.org/research-and-data/publications/business-review/1998/july-august/brja98dc.pdf.

Dai, Qiang and Singleton, Kenneth J. "Specification Analysis of the Affine Term Structure Models." *Journal of Finance*, October 2000, *55*(5), pp. 1943-978.

D'Amico, Stefania; Kim, Don H. and Wei, Min. "Tips from TIPS: The Informational Content of Treasury Inflation-Protected Security Prices." BIS Working Paper No. 248, Bank for International Settlements, 2008.

Duffee, Gregory R. "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance*, February 2002, *57*(1), pp. 405-43.

Duffee, Gregory R. "Are Variations in Term Premia Related to the Macroeconomy?" Working paper, Hass School of Business, University of California–Berkeley, March 2006.

Duffie, Darrell. *Dynamic Asset Pricing Theory*. Third Edition. Princeton, NJ: Princeton University Press, 2001.

Evans, Martin. "Discovering the Link Between Inflation Rates and Inflation Uncertainty." *Journal of Money, Credit, and Banking*, May 1991, *23*(2), pp. 169-84.

Hördahl, Peter; Tristani, Oreste and Vestin, David. "A Joint Econometric Model of Macroeconomic and Term-Structure Dynamics." *Journal of Econometrics*, March/April 2006, *131*(1-2), pp. 405-44.

Kim, Don H. "Challenges in Macro-Finance Modeling." Working Paper 2008-06, Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series, February 2008.

Kim, Don H. and Orphanides, Athanasios. "Term Structure Estimation with Survey Data on Interest Rate Forecasts." Working Paper 2005-48, Finance and Economics Discussion Series, Federal Reserve Board, October 2005; www.federalreserve.gov/pubs/feds/2005/200548/200548pap.pdf.

Kim, Don H. and Orphanides, Athanasios. "Bond Market Term Premium: What Is It, and How Can We Measure It?" *BIS Quarterly Review*, June 2007, pp. 27-40.

Kohn, Donald L. Comment on "Forward-Looking Rules for Monetary Policy," in John B. Taylor, ed., *Monetary Policy Rules*. Chicago: University of Chicago Press, 1999, pp. 157-202.

Kohn, Donald L. "Inflation Modeling: A Policymaker's Perspective." Remarks to the International Research Forum on Monetary Policy Conference, Frankfurt am Main, Germany, May 20, 2005; www.federalreserve.gov/Boarddocs/Speeches/2005/20050520/default.htm.

Litterman, Robert B. "Forecasting with Bayesian Vector Autoregressions—Five Years of Experience." *Journal of Business and Economic Statistics*, January 1986, *4*(1), pp. 25-38.

Litterman, Robert B. and Scheinkman, Jose A. "Common Factors Affecting Bond Returns." *Journal of Fixed Income*, June 1991, *1*(1), pp. 54-61.

Moench, E. "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach." *Journal of Econometrics*, September 2008, *146*(1), pp. 26-43.

Muth, John, F. "Optimal Properties of the Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 1960, *55*, pp. 299-306.

Rudd, Jeremy and Whelan, Karl. "Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?" *American Economic Review*, March 2006, *96*(1), pp. 303-20.

Rudebusch, Glenn, D. "Do Measures of Monetary Policy in a VAR Make Sense?" *International Economic Review*, November 1998, *39*(4), pp. 907-31.

Rudebusch, Glenn D. and Wu, Tao. "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy." Working Paper 2003-17, Federal Reserve Bank of San Francisco, 2003; www.frbsf.org/publications/economics/papers/2003/wp03-17bk.pdf.

Sangvinatsos, Antonios and Wachter, Jessica A. "Does the Failure of the Expectations Hypothesis Matter for Long-Term Investors?" *Journal of Finance*, February 2005, *60*(1), pp. 179-230.

Sims, Christopher A. "Macroeconomics and Reality." *Econometrica*, 1980, *48*, pp. 1-48.

Sims, Christopher A. Commentary on "Do Measures of Monetary Policy in a VAR Make Sense?" *International Economic Review*, 1998, *39*(4), pp. 933-41.

Stock, James H. and Watson, Mark W. "Evidence on Structural Instability of Macroeconomic Time Series Relations." *Journal of Business Economics and Statistics*, January 1996, *14*(1), pp. 11-30.

Stock, James H. and Watson, Mark W. "Vector Autoregressions." *Journal of Economic Perspectives*, Fall 2001, *15*(4), pp. 101-15.

Stock, James H. and Watson, Mark W. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, September 2003, *41*(3), pp. 788-829.

Stock, James H. and Watson, Mark W. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit, and Banking*, February 2007, *39*(s1), pp. 3-33.

Todd, Richard M. "Improving Economic Forecasting with Bayesian Vector Autoregression." Federal Reserve Bank of Minneapolis *Quarterly Review*, Fall 1984, *8*(4), pp. 18-29; www.minneapolisfed.org/research/QR/QR843.pdf.