Understanding the term structure of yield curve volatility

Anna Cieslak and Pavol Povala*

ABSTRACT

We study the joint behavior of the yield and volatility curves in the US Treasury market. Using almost two decades of high-frequency bond data, we obtain a so far unexplored view of the comovement between the yield curve states, structure of volatilities and their interactions with interest rates themselves. Based on that insight, we design a tractable model able to explain and decompose the dynamics of the two curves. The empirical success of our approach hinges upon two elements: a multivariate volatility whose sources of shocks and persistence are detached from those spanning yields, and the identification of volatility states with the support of realized covolatility proxies and filtering. The model separates three economically distinct elements of yield volatility: (*i*) a more erratic short-end state, (*ii*) a smoother long-end state, and (*iii*) a covolatility component capturing interactions between the long and the intermediate region of the curve. We show that the model-implied factors carry an interpretation and feature different responses to the economic environment. Expectations and uncertainty about the future path of key macro aggregates—rather than their realized numbers—can explain up to 95% of the latent factor variation. Interest rate volatility provides new information about the economic landscape that cannot be learned from observing yields at infrequent intervals. We find that different volatility states efficiently reflect the duration structure of uncertainties in the economy.

JEL classification: E43, C51
Keywords: interest rate risk, realized yield covolatility, affine models, conditional PCA, macro surveys
First version: May, 2009
This version: January, 2010

*Anna Cieslak and Pavol Povala are at the University of Lugano, member of the Swiss Finance Institute. Cieslak is visiting at the ChicagoBooth. Cieslak (corresponding author): acieslak@chicagobooth.edu, University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, IL 60637 Chicago. Povala: pavol.povala@lu.unisi.ch, University of Lugano, Institute of Finance, Via Buffi 13a, 6900 Lugano, Switzerland. We thank Torben Andersen, Luca Benzoni, John Cochrane, Jerome Detemple, Fabio Trojani, Pietro Veronesi and seminar participants at the Federal Reserve Bank of Chicago, University of Chicago, SFI NCCR PhD Workshop in Gerzensee for their comments. Our work has benefited from the access to the financial data warehouse and the computational platform (FDWH and FCP projects) sponsored by Prof. Fabio Trojani and developed by Tomasz Wisniewski at the University of Lugano. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK project under the direction of Prof. Trojani).

1. Introduction

With rich evidence collected on the evolution of the US interest rates, the yield volatility curve leaves several questions and controversies open. This paper studies the structure, economic content and pricing implications of the fluctuating covariance matrix of interest rates. Using almost two decades of tick-by-tick bond transaction data, we obtain a novel look into the development of interest rates and their volatilities, and show how these two pieces complement each other. This knowledge allows us to design a no-arbitrage term structure model able to simultaneously accommodate the dynamics of both curves.

We contribute to the existing yield curve literature along several dimensions. In contrast to previous studies that mainly focus on a simple volatility specification, we explain yield covolatilities across the whole spectrum of available maturities, and quantify their impact on the yield curve itself. The strong empirical fit and forecasting performance of our framework is just a first stage that stimulates a number of economic questions. Most importantly, we show that different volatility states efficiently encode the duration structure of uncertainties in the economy.

Our inquiry proceeds in several steps. In a model-free setup, we establish a set of empirical facts about the covariation of yields. We document strong and persistent movements in the conditional correlation between the level, slope and curvature that run counter the traditional logic of principal components. Although we do find a clear leading factor (most of the time the level), the portion of yield variation it explains can shift from 95% down to 50%. In those vibrant periods, other types of moves may well become more important. As such, static decompositions of interest rates into the level, slope and curvature fail to capture the full extent of the interest rates dynamics. The covariance matrix of yields does vary over time, and we uncover that at least three different forces are at work in determining those moves. While state variables generating volatilities appear largely detached from those generating yields—a phenomenon which has earned the label of the unspanned stochastic volatility (USV)[1]—we find an asymmetric coincidence between the level of interest rates and volatilities: The Fed's moves correspond to periods of elevated rate volatility, and its shift is more pronounced during cuts than it is during hikes.

Such combination of features presents a challenge for those seeking to explain interest rates jointly with their second moments. The assumption underlying many term structure models—that three factors are enough—appears unsuited in context of this goal. While three factors suffice for modeling interest rates alone, additional flexibility is required to capture volatilities. We exploit this intuition when translating empirical results into the model. First, the model has three standard yield curve factors. On top of it, to generate an adequate variation in the covariance matrix of yields, we introduce a dynamic dependence between those factors. The stochastic covariance process we adopt gives rise to a three-variate model of yield volatilities, and thus aligns well with the empirical properties of the realized yield covariation. Second, we are careful about the weak relationship between yields and volatilities. The USV models impose their separation by an explicit parametric restriction. Our approach is milder. We take as a clue the simple observation that yields and volatilities are governed by distinct statistical rules. Thus, it is unlikely that the same variable

[1] See Collin-Dufresne and Goldstein (2002) for the origins of this expression. We survey this work in the literature review section below.

plays an equal role in determining both. Our specification discriminates between the sources of persistence and shocks in yield factors from those in volatilities. With sufficient flexibility along both dimensions, no additional parameter constraints turn out to be required to explain the data.

While few would contend that higher-dimensional settings are key to explain the interest rate risk (e.g. Joslin, 2007; Kim, 2007b; Andersen and Benzoni, 2009), the most comprehensive models typically do not exceed dimension four and allow at most two volatility states. Going beyond this scope appears empirically desirable, yet it also raises two important concerns: (i) high parametrization, and (ii) the inability to identify bond volatilities from yields alone.[2] The structure we consider here adds flexibility for modeling volatility, but with 13 parameters in the physical dynamics it remains tractable by the standards of affine term structure models (ATSMs). As for the second concern, we do agree that identification of volatilities from the low-frequency yield data could frustrate any model. However, by constructing a close proxy for the stochastic covariance matrix of yields, and exploiting it in estimation, we are not troubled by this point.

The model achieves good performance in explaining the yield covariance matrix across a range of maturities without sacrifices in fitting yields. Its statistical record authorizes our investigation of the model-implied states, which so far has not been ventured in the literature. Several findings are worth highlighting.

First, the identified volatility states correspond closely with the roles of the respective yield curve factors. In both, we disentangle short- versus long-end components that play distinct roles along the curve. Those shorter-term factors, governing yields and volatilities at maturity of two to three years, are more transient and erratic. The longer-term ones, instead, that are responsible for maturities from beyond five years, exhibit more persistent and smoother dynamics. Additionally, we identify a covolatility state, which captures the comovement between the long and the intermediate region of the curve. The presence of several volatility variables extends the evidence pertaining to the preferred affine models, $A_1(n)$,[3] which have dominantly focused on the single factor short-rate volatility.

Second, while we do not exclude volatility states from entering the cross-section of yields, we find their cross-sectional importance to be quantitatively negligible compared to that of the interest rate states. In fact, only the long-end volatility factor can move the yield curve by a visible amount. But its impact does not exceed seven basis points on average, and is limited to maturities above five years. There is little hope that the information about volatilities can be extracted from the cross-section of yields alone. We show that the backing-out approach, i.e. inverting the term structure equation from a subset of yields to recover states, fails at identifying even one, let alone three volatility factors. For that reason, our estimation relies on filtering and extra spot covolatility measurements. This latter choice sets our evidence apart from the literature relying on interest rate derivatives that, in general, are not delta-neutral, and involve an additional

[2] Joslin (2006, 2007) discusses a model with two conditionally Gaussian and two CIR states, one of which is dedicated to volatilities. The volatility state is identified from interest rate derivatives. By Dai and Singleton (2000) such a model is classified as $A_2(4)$. Outside the pure latent factor domain, macro-finance delivers some examples of models with more than two variables displaying stochastic volatility. We discuss them in the literature review below. An important group of papers studies high dimensional HJM settings (e.g. Trolle and Schwartz, 2009; Han, 2007). Yet, by taking the current yield curve as given, these are not of direct comparison to the equilibrium motivated ATSMs.

[3] The naming convention used here follows the convenient taxonomy of ATSMs introduced by Dai and Singleton (2000). $A_m(n)$ denotes an n-factor model, in which m factors feature stochastic volatility.

layer of modeling assumptions. The use of realized covolatility, instead, gives us a clean view of the volatility factors, and allows a precise assessment of their interactions with the yield curve.

Last but not least, we find that the model-implied factors contain economic information, even though they are not directly linked to specific macro quantities by design. Given the notion that prices should reflect economic prospects more than past events, we use survey-based forecasts about key macro variables instead of their realized numbers. The combination of expectation and uncertainty proxies turns out to be highly informative about the filtered dynamics, being able to explain up to above 90% and 50% of variation in the yield and volatility factors, respectively. Importantly, macro variables related to short- versus long-run states form disjoint sets. Stemming from long-duration bonds, the longer-end volatility shows a pronounced response to the persistent real activity measures such as expectations of GDP growth, or uncertainty about unemployment. The short-term volatility, in turn, is linked to the uncertainty and expectations about the monetary policy, and uncertainties surrounding inflation and industrial production, both of which give a more short-lived description of the economic conditions in our sample. Prospects of the housing sector emerge as the only variable with a jointly significant effect on short- and long-end components. Finally, the covolatility state is associated with uncertainty proxies on monetary policy and the real economy, showing that those variables influence the correlation between the intermediate and long region of the curve.

Our work evolves around the central theme that interest rates and their volatilities live separate lives. Intuitively, the yield curve at any point in time represents a set of equilibrium prices of bonds—or equivalently risk-adjusted consensus expectation about the future path of interest rates—after market forces have been at work. Interest rate volatility, in contrast, tracks mainly the process of finding those equilibrium expectations over time. An increased uncertainty about the real and nominal sectors or monetary policy usually makes it more challenging for the market participants to find equilibrium prices which, in turn, boosts the volatility. In this way, interest rate volatility captures a complementary piece of information that cannot be learned from observing yields at infrequent intervals.

Recent term structure environment gives a good illustration of the complexity of the two objects. In Table 1, we gather major moves in the yield and volatility curves from 2000 to 2004, and provide snapshots of the underlying economic circumstances. This interesting period covers the longest easing cycle in our sample, during which the Fed brought interest rates down to a (then) unprecedented level of just 1%. When the uncertainty about economic outcomes is high, the volatility curve shows motion. Indeed, at least part of the variation in interest rate risk should be associated with the varying perceptions about key macro variables (e.g. Kim, 2007a). Since our volatility measures arise from active trading in liquid bonds, we would expect them to give a timely reflection of changing market expectations and fears. Fed officials watch this uncertainty to take their interest rate decisions. The transcripts from the FOMC meetings provide a good record of this point, e.g. "Uncertainty about key interactions in the economy is a good reason to wait" (Mr. Donald Kohn, October 3, 2000).[4] Our event Table 1 shows that the volatility curve can be at least as rich a source of information as yields themselves. In this paper, we take an attempt towards its comprehensive modeling and thus hope to contribute to a more complete understanding of the behavior of interest rates.

[4]See the FOMC transcript at http://federalreserve.gov/monetarypolicy/files/FOMC20001003meeting.pdf.

Related literature

Recent research into interest rate volatility has evolved in at least three loosely related directions. Below, we provide a review of their different leading themes: (*i*) unspanned factors, (*ii*) realized volatility and jumps, (*iii*) non-Gaussian models. Our work touches upon these three strands.

Unspanned volatility. As a prediction of ATSMs, bond markets are complete. Empirically, however, bond market incompleteness tends to reveal itself through a poor performance of bond portfolios in hedging interest rate derivatives, and more generally, in hedging interest rate volatility risk. Several papers document a weak relation between the bond volatility, realized as well as derivative-based, and the spot yield curve factors. Collin-Dufresne and Goldstein (2002, CDG), Heidari and Wu (2003), and Li and Zhao (2006), among others, all conclude that movements in yields explain at best only about a half of the variation in interest rate derivatives.

Not surprisingly, the hypothesis of such unspanned risk in the bond market has provided for both active research and a controversy in the affine term structure literature. CDG (2002) formalize its intuition for the standard three-factor ATSMs under the name of unspanned stochastic volatility. The USV label reveals the only possible source of incompleteness in ATSMs: If anything, unspanned must be the volatility. CDG (2002) characterize parameter restrictions such that the instantaneous interest rate volatility does not affect the cross-section of yields.

By providing testable predictions, the USV theory has triggered increased interest in the ATSMs' ability to explain the yield volatility dynamics. Here the evidence is mixed. In support of the unspanning hypothesis, Collin-Dufresne, Goldstein, and Jones (2008, CDGJ) report that, over the 1988–2005 sample, variance series generated by the standard $A_1(3)$ model are essentially unrelated to the model-free conditional volatility measures. Jacobs and Karoui (2009) find a correlation up to 75% using the same model estimated on Treasury yields over the 1970–2003 period. In the more recent sample 1991–2003, however, this correlation breaks down and becomes slightly negative at the long end of the curve.

While imposing the USV restriction does improve the ability of low-dimensional ATSMs to fit the time-series of volatilities, it also comes at a cost of higher cross-sectional pricing errors. As a consequence, several papers reject the USV in favor of an unconstrained model (Bikbov and Chernov, 2009; Joslin, 2007; Thompson, 2008). This evidence mostly concerns ATSMs with four factors[5] and a univariate volatility structure. What lies at the heart of the unspanning hypothesis, however, is an economic effect more than a failure of a specific model restriction. In line with this intuition, Kim (2007b) highlights a major demand for term structure models, not necessarily USV ones, with sufficient flexibility to match jointly yields and their volatility dynamics. Our analysis focuses on designing, implementing, and deriving the implications of such a model. For the 1992–2007 sample period, the model-generated term structure of conditional volatilities consistently tracks the observed series with an R^2 exceeding 96%.

[5] Bikbov and Chernov (2009) test the USV in the $A_1(3)$ model studied by CDG (2002). Thompson (2008) extends the evidence to the $A_1(4)$ USV model. $A_2(4)$ is considered in Joslin (2007).

Realized volatility. In combination with recent advances in high-frequency finance, the USV debate has encouraged a new model-free look into the statistical properties of bond volatility. Andersen and Benzoni (2009, AB) test empirically the linear spanning restriction of ATSMs using measures of realized volatility over the 1991–2000 period. They confirm that systematic volatility factors are largely independent from the cross section of yields, and call for essential extensions of the popular models on the volatility front. While our model is cast within the general affine framework, we attain this goal by combining two ingredients: the rich form of covolatility states plus their identification from the realized data.

The availability of high-frequency observations from spot and futures fixed income markets has revived interest in the impact of economic news releases on bond return volatility and jumps. As such, this literature has remained mostly empirical. Across studies, a consensus view is that a large fraction (about 70%) of big moves in bond prices and thus in realized volatilities occurs on pre-scheduled macroeconomic announcements (e.g., Beechey and Wright, 2008; Dungey, McKenzie, and Smith, 2009; Jiang, Lo, and Verdelhan, 2009).[6] Focusing on the discontinuities in bond returns, Wright and Zhou (2009) show for instance that the mean jump size extracted from the 30-year interest rate futures has a significant predictive power for expected excess bond returns. Indeed, relative to other liquid asset markets, bond prices tend to provide the most clear and pronounced reaction to economic news (Andersen, Bollerslev, Diebold, and Vega, 2007; Jones, Lamont, and Lumsdaine, 1998). These studies suggest that a rich economic content is present in bond volatilities. We provide a model-based decomposition of the volatility curve, and find that its components have different reactions to measures of economic conditions.

From latent to macro-motivated models. Our knowledge of the links between interest rate volatility and the macroeconomy coming from no-arbitrage yield curve models is still sparse. Most of the macro-finance term structure literature, cast in a Gaussian-affine framework, has by construction remained silent about the yield volatility.[7]

Several recent papers mark an important development by going beyond the standard Gaussian setup. Examples include Campbell, Sunderam, and Viceira (2009, CSV), Adrian and Wu (2009), Hautsch and Ou (2008), Bekker and Bouwman (2009) or Haubrich, Pennacchi, and Ritchken (2008, HPR). Some features unify the economics of these models. In particular, volatility is multivariate, and reflects different sources of risk. Both, Adrian and Wu (2009) and CSV highlight the importance of stochastic covariation between the real pricing kernel and (expected) inflation in determining excess bond returns. Likewise, Hautsch and Ou (2008) find ex post that the extracted persistent volatility factors are important for explaining bond excess returns. HPR introduce a similar effect through the relationship between inflation and the real interest rate.

In an extension of the latent factor approach, these models attach economic labels to different yield volatility components. To the extent that the volatility itself remains unobservable or is extracted from an

[6]The interpretation of these results is limited to short spans of data. For Treasury bonds, the longest sample is considered in Andersen and Benzoni (2009), who use the GovPX database covering years 1991 through 2000. Samples used in other studies typically do not exceed four years. For instance, the analysis in Jiang, Lo, and Verdelhan (2009) is limited to a two-year period of 2005–2006 marking a particular time, in which bond market volatility stroke historically lowest level. Longer samples are considered in studies involving interest rate futures.

[7]The literature using Gaussian-affine no arbitrage setting in conjunction with macro factors is vast. Notable examples of such models are Ang and Piazzesi (2003), Ang, Dong, and Piazzesi (2007), Bikbov and Chernov (2008), Duffee (2006, 2007).

auxiliary model, the identification and interpretation of its components relies on specific model assumptions. Indeed, explaining the volatility curve per se is not in direct focus of those models. We, in contrast, start completely latent, and having explained yields and volatilities, try to understand the impact of economic quantities on the states forming both curves.

2. Empirical evidence

This section describes our data set. We discuss the properties of yield covolatilities and their relation with the yield curve. Evidence collected here sets the stage for our model design in Section 3.

2.1. Data

This paper is a first attempt to analyze yield volatility with the help of high-frequency Treasury bond data spanning two long expansions, one recession and three monetary cycles in the US economy. We obtain 16 years' worth of high-frequency price data of US Treasury securities covering the period from January 1992 through December 2007. We construct the sample by splicing historical observations from two inter-dealer broker (IDB) platforms: GovPX (1992:01–2000:12) and BrokerTec (2001:01–2007:12). The merged data set covers the majority of transactions in the US Treasury secondary market with a market share of 60% and 61% for GovPX and BrokerTec, respectively (Mizrach and Neely, 2006). As such, our dataset provides a comprehensive description of the contemporary yield curve environment. Beside the unavailability of high frequency Treasury bond data prior to 1991 when the GovPX started operating, other reasons speak against considering longer samples. Most importantly, there is empirical evidence that the conduct of monetary policy changed significantly during the eighties (e.g., Ang, Boivin, Dong, and Loo-Kung, 2009). The market functioning has also shifted dramatically with the advent of computers (automated trading), interest rate derivative instruments, and the swap market. Capturing such institutional features is not the object of our analysis.

GovPX comprises Treasury bills and bonds of maturities: three, six and 12 months, and two, three, five, seven, ten and 30 years. BrokerTec, instead, contains only Treasury bonds with maturities: two, three, five, ten and 30 years. In the GovPX period, we identify on-the-run securities and use their mid-quotes for further analysis. Unlike GovPX, which is a voice-assisted brokerage system, BrokerTec is a fully electronic trading platform attracting vast liquidity and thus allowing us to consider traded prices of the on-the-run securities. In total, we work with around 37.7 million on-the-run Treasury bond quotes/transactions. Appendix A.1 reports the average number of quotes and trades per day that underlie our subsequent analysis.

The US Treasury market is open around the clock, but the trading volumes and volatility are concentrated during the New York trading hours. Roughly 95% of trading occurs between 7:30AM and 5:00PM EST (see also Fleming, 1997). This interval covers all major macroeconomic and monetary policy announcements, which are commonly scheduled either for 9:00AM EST or 2:15PM EST. We consider this time span as a trading day. Especially around US bank holidays, there are trading days with a very low level of trading activity. In such cases, we follow the approach of Andersen and Benzoni (2009) and delete days with no trading for more than three hours.

We sample bond prices at ten-minute intervals taking the last available price for each sampling point. We choose this sampling frequency so that it strikes the balance between the non-synchronicity in trading and the efficiency of the realized volatility estimators (Zhang, Mykland, and Ait-Sahalia, 2005). The microstructure noise does not appear to be an issue in our data, as indicated by the volatility signature plots and very low autocorrelation of equally spaced yield changes (see Appendix A.2).

While the raw data set contains coupon bonds, it is crucial for our analysis to have precise and timely estimates of zero coupon yields. Using equally-spaced high-frequency price data, we construct the zero coupon yield curve for every sampling point. To this end, we apply smoothing splines with roughness penalty as described in Fisher, Nychka, and Zervos (1994). We purposely avoid using the Nelson-Siegel type of method because it seems to wash out some valuable information, as reported by Cochrane and Piazzesi (2008). The technical details on our zero coupon yield curve estimation are collected in Appendix A.3.

The liquidity in the secondary bond market is concentrated in two-, three-, five- and ten-year securities (see also Fleming and Mizrach, 2008, Table 1). We assume that the dynamics of this most liquid segment spans the information content of the whole curve. Since any bootstrapping method is precise for maturities close to the observed yields, for subsequent covolatility analysis we select yields which are closest to the observed coupon bond maturities.

2.2. Realized yield covariances

Our analysis of interest rate risks focuses on nominal bonds. The high-frequency zero curve serves as an input for the calculation of the realized covariance matrix of yields. We consider zero yields with two, three, five, seven and ten-year maturities. Let y_t be the vector of yields with different maturities observed at time t. Time is measured in daily units. The realized covariance matrix is constructed by summing up outer products of a vector of ten-minute yield changes, and aggregating them over the interval of one day $[t, t+1]$:

$$RCov(t, t+1; N) = \sum_{i=1,\ldots,N} \left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}} \right) \left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}} \right)'. \tag{1}$$

$N = 58$ is the number of equally spaced bond prices (yields) per day t implied by the ten-minute sampling, and i denotes the i-th change during the day. The weekly or monthly realized covariances follow by aggregating the daily measure over the corresponding time interval. To obtain annualized numbers, we multiply $RCov$ by 250 for daily, 52 for weekly or 12 for monthly frequency, respectively. Based on Jacod (1994) and Barndorff-Nielsen and Shephard (2004), for frequent sampling the quantity (1) converges to the underlying quadratic covariation of yields. In Section 7, we positively assess the robustness of this estimator, and compare it to the alternatives proposed in the literature.

We aim to ensure that our volatility measures reflect views of active market participants rather than institutional effects. This motivates the following two choices: First, our construction of $RCov$ dynamics relies exclusively on the within-day observations, thus excluding the volatility patterns outside the US trading hours. Even though between- and within-day volatilities track each other fairly closely, we observe several instances of substantial differences and abrupt spikes in the between-day volatility pattern, which we cannot

relate to any major news on the US market. To account for the total magnitude of volatility, we add to the within-day number the squared overnight yield change from close (5:00PM) to open (7:30PM). We then compute the unconditional average of the total and within-day realized yield covariation, respectively, and each day scale the within-day *RCov* dynamics by the total-to-within ratio.

The second choice lies in focusing on the intermediate and long maturities (two to ten years), i.e. very liquid and frequently traded bonds. The short end of the curve (maturities of one year and below) is deliberately excluded from the realized covariance matrix computations for several reasons. Over our sample period, this segment of the curve exhibited a continuing decline in trading and quoting activity, and was completely suspended in March 2001 (see Appendix A.1).[8] A lower liquidity at the short end of the on-the-run curve is documented by Fleming (2003). Moreover, relative to the latter part, the dynamics of the short segment is complicated by its interactions with the LIBOR market and monetary policy operations. Such distortions, while interesting in their own right, are not directly relevant to the analysis we perform.

2.3. Information in second moments of yields

Table 2 reports summary statistics for weekly yields (panel *a*) and realized volatilities (panel *b*). Fig. 1 plots average curves, both unconditional and contingent upon the monetary policy cycle. A monotonically increasing term structure of average yields is accompanied by a humped term structure of volatilities, with the hump occurring at the three-year maturity. In our sample, monetary easing not only increases the slope of the first curve, but also lifts the level in the latter and the magnitude of the hump. While both exhibit non-normalities, the statistical properties of the two objects are very different. Not surprisingly, the non-normality becomes more pronounced in the term structure of volatilities. Compared to the smooth evolution of yields, the volatility curve experiences periods of elevated and abruptly changing dynamics apparent in Fig. 2. This autonomy of the volatility process encourages a more detailed look into its behavior. In the remainder of this section we examine the second moments of yields along three dimensions: (*i*) changes in their dynamic properties, (*ii*) the number of underlying factors, and (*iii*) their potential link to the yield levels.

[Fig. 2 about here]

Level, slope and curvature viewed dynamically. Much of the intuition about factors driving the zero curve has been obtained from the principal component analysis (PCA) of the unconditional covariances of yields (Litterman and Scheinkman, 1991). Such analysis aggregates decades of yield curve information into a single set of numbers. In contrast, Eq. (1) combined with the high-frequency data provides a proxy for the unobservable conditional covariance matrix, and allows its dynamic decomposition. This step tells us that the unconditional PCA, usually applied to motivate three-factor models, washes away some valuable information about factors driving yields. While dynamically we do find three main factors, their relative importance fluctuates over time. The nature of factors can change with instances of slope moves taking the

[8] The decline in the trading of short maturity bonds is not particular to GovPX or BrokerTec data. A similar development took place in the interest rate futures market.

lead over level moves, and curvature moves—over the slope. The portion of yield variation explained by the level factor, typically exceeding 90%, can at times drop to just above 50%.[9]

If the standard PCA gives an incomplete description of yields, this shall be revealed in the time-varying comovement of factors it implies. To investigate this channel, we decompose the unconditional covariance matrix of four yields with maturities two, three, five and ten years. Factor loadings from the unconditional PCA serve to construct the level, slope and curvature tick-by-tick, and to estimate their realized correlations as plotted in Fig. 3.[10] All correlations display a persistent pattern over time. For instance, the correlation between the slope and the level factor oscillates between ±50%, and is generally lower during periods of monetary easings. Intuitively, interest rate cuts tend to increase the slope of the curve on concerns of looming inflation. Instead, periods of monetary stability make the short and long end of the curve more independent. However, the superposition of these correlations against the Fed regimes in Fig. 3 also shows that interactions between factors are more complex than just a monetary policy response. For the modeling, this picture translates into the requirement of time-varying dependence between state variables determining yields.

[Fig. 3 about here]

Factors in volatilities. The time-varying nature of loadings in the conditional PCA suggests that multiple stochastic volatility factors act on the second moments of yields. We find that yield volatilities do not move on a single determinant. Similar to the cross-section of yields, at least three factors are also needed to explain the dynamics of the realized volatility curve, which exhibits the familiar level, bent slope and curvature-like movements (panel *a*, Fig. 4). The first three principal components explain 90.2%, 6.1% and 2.3% of its variation (panel *b*). Importantly, this observation comes from analyzing middle to long yields only, and thus is not driven by idiosyncratic volatility at the very short end of the curve.

[Fig. 4 about here]

From the modeling perspective, this result raises a natural question: Are factors driving volatilities related to those those typically found in yields?

Link between interest rates and volatilities. Much of the theoretical and empirical evidence points to a link between the level of interest rates and their volatility. The affine or quadratic models, for instance,

[9] Details of the dynamic decomposition are omitted to conserve space and are available upon request. Additionally, we perform a formal log-likelihood test (see e.g. Fengler, Härdle, and Villa (2001) for details of the testing procedure). We find a strong rejection of: (*i*) a constant covariance matrix hypothesis (constant eigenvalues and eigenvectors), and (*ii*) a common principal component hypothesis (constant eigenvectors but time-varying eigenvalues). Both hypotheses are tested against the alternative under which the covariance matrices do not have a constant factor structure across subsamples. In addition to performing the test over the whole sample period, we check the stability of monthly conditional covariances year-by-year. The test consistently rejects both hypotheses. While the rejection of the first one is not surprising and could easily arise from changing yield volatilities, the latter one is more important: It confirms that the space spanned by eigenvectors, giving rise to the level, slope and curvature interpretation, is in fact not stable across periods.

[10] The results persist if we follow a different portfolio construction that is immunized to all but one type of yield curve movements, and rebalanced at the beginning of each month. The dynamics of factor correlations are very close to those obtained with the unconditional PCA loadings.

imply that the same subset of factors determines both yields and their volatilities. As one simple example, a single-factor CIR model suggests that the volatility is high whenever the short rate is high—a prediction that remained valid through the early 1980s (Chan, Karolyi, Longstaff, and Sanders, 1992).

More recently, however, the USV literature has argued that the yield-volatility relation is in fact weak. We add to this evidence by showing their more complex interactions than those implied by the linear regressions used in the USV tests. Fig. 5 scatter-plots weekly realized volatilities against the level of interest rates with a matching maturity. The shape of the nonparametric regression fitted to the data discards the possibility of a positive correlation between yields and volatilities in our sample period. If any, the relationship appears to be asymmetrically U-shaped, which clearly contrasts with the early 1980s' episode. The volatility is low for the intermediate interest rates range, and increases when rates move to either end of the spectrum. The rise in volatility is more pronounced in low interest rate regimes, and thus explains the negative unconditional correlations between yields and volatilities reported in panel *c* of Table 2. The last panel of Fig. 5 provides a simple illustration of this point: Federal funds rate cuts induce a stronger upward revision in volatility than do tightenings. The asymmetry is most pronounced for shorter maturities (two years) and decays at the longer end of the curve.[11]

Not surprisingly, evidence on the relationship between interest rates and volatility is mixed and controversial. Fig. 6 provides one explanation to the lacking consensus. We plot beta-coefficients and R^2's in regressions of realized volatilities on yields with two-, five- and ten-year maturities performed on daily data over a four-month rolling window. The yield-volatility link turns out to be highly state dependent. It reveals large fluctuations in the R^2 and switching signs of the regression coefficients. Yet, despite the obvious instability, it is hard to argue that the link is completely non-existing and should be discarded as a matter of principle.

[Fig. 5 and Fig. 6 about here]

For the sake of model design, we can learn how shocks in yields and volatilities are interrelated by estimating a VAR for the joint system. We include three bond portfolios mimicking the level, slope and curvature of the yield curve plus three realized volatility factors: the level RV_t^{2Y}, the slope $(RV_t^{10Y} - RV_t^{2Y})$ and the covariance $RCov_t^{5Y,10Y}$. The highest cross-correlation between shocks is 16% (the yield level portfolio and the volatility level). This low correlation is economically intuitive: To the degree that the volatility is related to the process of finding equilibrium, its shocks should not be correlated with those in the equilibrium process (the yield curve).

[11] To understand the type of non-linearity in the yield-volatility relationship, we fit a generalized additive model (GAM) with a linear and a spline part, i.e. $v_t^\tau = \beta_0 + \beta_1 y_t^\tau + s(y_t^\tau) + \varepsilon_t$. The result of the exercise is twofold. For short and intermediate maturities (of two and five years) both components are significant, with a negative β_1 loading and an asymmetric spline component. For the long end of the curve (ten years), in turn, we find no support of a linear component, and a weak confirmation of the U-shaped asymmetry.

3. The model

Evidence of the previous section provides guidelines for our modeling approach. First, to generate a sufficiently rich variation in covariances of yields, we allow a multivariate volatility and dynamic interactions between factors. Second, while our empirical findings do invoke the notion of unspanning, we remain cautious about imposing it within the model. In fact, we show that such restriction is not required to fit yields and volatilities jointly. Rather, given findings on the number of factors in yields and volatilities, we do not expect a low-dimensional model (whether or not USV) to perform well on both fronts. In the remainder of this section, we formulate a sufficiently flexible model, and later verify its viability in terms of the econometric fit and economic interpretation of factors.

Our benchmark model is cast in a reduced-form continuous-time framework. In specifying the state dynamics, we take an agnostic view on factor labels, but assign them to two groups: (i) expectations factors X_t, and (ii) covariance factors V_t. The physical dynamics are given by the system:

$$dX_t = (\mu_X + \mathcal{K}_X X_t)dt + \sqrt{V_t}dZ^{\mathbb{P}}_{X,t} \tag{2}$$

$$dV_t = (\Omega\Omega' + MV_t + V_t M')\,dt + \sqrt{V_t}dW^{\mathbb{P}}_t Q + Q' dW^{\mathbb{P}\prime}_t \sqrt{V_t}, \tag{3}$$

where X_t is a n-vector, and V_t is a $n \times n$ process of symmetric positive definite matrices—a covariance matrix process proposed by Bru (1991) and studied by Gourieroux, Jasiak, and Sufana (2005). Accordingly, $Z^{\mathbb{P}}_X$ and $W^{\mathbb{P}}$ are a n-dimensional vector and a $n \times n$ matrix of independent Brownian motions.[12] μ_X is a n-vector of parameters and $\mathcal{K}_X, M$ and Q are given as $n \times n$ parameter matrices. To ensure a valid covariance matrix process V_t, we specify $\Omega\Omega' = kQ'Q$ with an integer degrees of freedom parameter k such that $k > n-1$, and require that Q is invertible. This last condition guarantees that V_t stays in the positive definite domain (see e.g., Gourieroux, 2006).

The short interest rate is an affine function of X_t variables, but contains an additional source of persistent shocks:

$$r_t = \gamma_0 + \gamma'_X X_t + \gamma_f f_t. \tag{4}$$

The state f_t evolves as:

$$df_t = (\mu_f + \mathcal{K}_f f_t + \mathcal{K}_{fX} X_t)dt + \sigma_f dZ^{\mathbb{P}}_{f,t}, \tag{5}$$

with $Z^{\mathbb{P}}_{f,t}$ denoting a single Brownian motion independent of all other shocks in the economy. $\gamma_f, \mathcal{K}_f$ and σ_f are scalars, and γ'_X and $\mathcal{K}_{fX}$ are $(1 \times n)$-vectors of parameters. For convenience, we collect X_t and f_t factors in a vector $Y_t = (X'_t, f_t)'$, whose dynamics can be compactly expressed as:

$$dY_t = (\mu_Y + \mathcal{K}_Y Y_t)\,dt + \Sigma_Y(V_t)dZ^{\mathbb{P}}_t, \tag{6}$$

[12] It is straightforward to introduce correlations between Z_X and W by setting $dZ = dW\rho + \sqrt{1-\rho'\rho}dB$ for some constant vector ρ, where dB is an n-dimensional Brownian motion independent of columns in dW. We state the general solution for $\rho \neq 0$ in the Appendix C. However, based on empirical findings of Section 2, in particular the very low correlation of shocks between volatilities and yields, we believe that this extension has no merit for the problem at hand.

with a block diagonal matrix $\Sigma_Y(V_t)\Sigma_Y(V_t)' = \begin{pmatrix} V_t & 0 \\ 0 & \sigma_f^2 \end{pmatrix}$ and $\mathcal{K}_Y = \begin{pmatrix} \mathcal{K}_X & 0_{n\times 1} \\ \mathcal{K}_{fX} & \mathcal{K}_f \end{pmatrix}$.

Bonds in this economy are priced using the standard no-arbitrage argument. By its convenience, we can abstract from a particular preference structure, and specify a general reduced-form compensation $\Lambda_{Y,t}$ required by investors to face shocks in the state vector:

$$\Lambda_{Y,t} = \Sigma_Y^{-1}(V_t)\left(\lambda_Y^0 + \lambda_Y^1 Y_t\right), \tag{7}$$

where λ_Y^0 is a $(n+1)$-vector and λ_Y^1 is a $(n+1)\times(n+1)$ matrix of parameters. To be viable, this formulation requires the invertibility of each block of matrix $\Sigma_Y(V_t)$, which is ensured by the positive-definiteness of $\sqrt{V}_t$, and with σ_f different from zero. In Eq. (7) we assume that only Z shocks are priced. Therefore, the risk neutral dynamics of X_t follow from the standard drift adjustment:

$$\mu_Y^{\mathbb{Q}} = \mu_Y - \lambda_Y^0 \tag{8}$$

$$\mathcal{K}_Y^{\mathbb{Q}} = \mathcal{K}_Y - \lambda_Y^1, \tag{9}$$

and the dynamics of V_t remain unchanged. It is technically possible to introduce a priced volatility risk without loosing the flexibility of the framework. However, in the presence of a weak spanning of volatility states by bonds it is difficult (if not impossible) to identify the market price of risk for volatility from bonds alone. For this, additional volatility-sensitive instruments such as bond options are needed.

Prices of nominal bonds are obtained by solving $P_t^\tau = E_t^{\mathbb{Q}}\left(e^{-\int_0^\tau r_s ds}\right)$. By the Feynman-Kac argument, and using the infinitesimal generator for the joint process $\{Y_t, V_t\}$, the solution for the nominal term structure has a simple affine form (see Appendix C):

$$P(t,\tau) = e^{A(\tau)+B(\tau)'Y_t+Tr[C(\tau)V_t]}, \tag{10}$$

where $Tr(\cdot)$ denotes the trace operator. The coefficients $A(\tau), B(\tau)$ and $C(\tau)$ solve a system of ordinary differential equations:

$$\frac{\partial A(\tau)}{\partial \tau} = B(\tau)'\mu_Y^{\mathbb{Q}} + \frac{1}{2}B_f^2(\tau)\sigma_f^2 + kTr\left[Q'QC(\tau)\right] - \gamma_0 \tag{11}$$

$$\frac{\partial B(\tau)}{\partial \tau} = K_Y^{\mathbb{Q}\prime}B(\tau) - \gamma_Y \tag{12}$$

$$\frac{\partial C(\tau)}{\partial \tau} = \frac{1}{2}B_X(\tau)B_X(\tau)' + C(\tau)M + M'C(\tau) + 2C(\tau)Q'QC(\tau), \tag{13}$$

where we split the $B(\tau)$ loadings as $B(\tau) = [B_X(\tau)', B_f(\tau)]'$ and $\gamma_Y = (\gamma_X', \gamma_f)'$. The boundary conditions for the system (11)–(13) are $A(0) = 0_{1\times 1}, B(0) = 0_{(n+1)\times 1}$ and $C(0) = 0_{n\times n}$. The $B(\tau)$ loadings have a simple form typical to the Gaussian models, and allow an immediate solution. The $C(\tau)$ matrix solves a matrix Riccati equation.

Defining $y_t^\tau = -\frac{1}{\tau}\ln P_t^\tau$, the term structure of interest rates has the form:

$$y_t^\tau = -\frac{A(\tau)}{\tau} - \frac{B(\tau)'}{\tau} Y_t - Tr\left[\frac{C(\tau)}{\tau} V_t\right]. \tag{14}$$

As a consequence of the dynamics (2)–(3), yields are an affine function of the entire state vector $(Y_t, \text{vec}(V_t)')'$. Under uncorrelated shocks dZ and dW, and the short rate (4), we leave only one channel open through which volatility states appear in the yield curve equation, i.e. the diffusive term in the Y_t dynamics (6). Thus, the instantaneous yield covariation is exclusively driven by the covariance factors:

$$\begin{aligned} v_t^{\tau_i,\tau_j} &:= \frac{1}{dt}\langle dy_t^{\tau_1}, dy_t^{\tau_2}\rangle \\ &= \frac{1}{\tau_1\tau_2}\{Tr\left[B_X(\tau_2)B_X(\tau_1)' + 4C(\tau_2)Q'QC(\tau_1)\right] V_t + B_f(\tau_1)B_f(\tau_2)\sigma_f^2\}. \end{aligned} \tag{15}$$

3.1. Discussion

In the basic setup, we consider three variables in Y_t, i.e. f_t plus a two-dimensional vector X_t. The latter is equipped with a 2×2 covariance matrix V_t. The form of V_t leads to a two-plus-one-variate process with two volatility plus a covariance factor. Thus, it presents a three-factor model of yield volatilities. The combination of six factors gives us a scope to fit both yields and their volatilities.

Even though we choose to follow the latent state tradition, the non-observability of factors does not preclude an attempt of their interpretation. We think of f_t as a short-term monetary policy factor. X_t, instead, represents longer-term forces that reflect expectations about the key elements of the economic landscape, e.g. the real and nominal sector. Naturally, they can impact the conditional expectation of f_t. It is through the longer-term factors that the time-varying volatility enters into the picture: V_t describes the amount of risk present in the economy, with its out-of-diagonal element V_{12} determining the conditional mix between X_t's.

Our split between volatility and expectation variables evokes the $A_m(n)$ classification of Dai and Singleton (2000). Still, at least two differences are worth highlighting. First, V_t represents a complete covariance matrix dynamics (i.e. volatilities plus covariances), and as such involves components which can switch sign. In contrast, independent CIR processes in ATSMs generate stochastic volatility one-by-one. Therefore, the covariances they imply are a linear combination of the volatility factors. Second, to make the roles of factors precise and interpretable, our specification intentionally excludes any interactions between V_t and X_t via the drift. Even though in ATSMs such interactions are usually allowed, we show in estimation that they are not called for by the data.

As dimensions grow, flexibility typically comes at the price of parsimony. The $A_1(4)$ version estimated by Joslin (2007) and Thompson (2008) involve over 20 parameters after excluding the market prices of risk. By these standards, the state space we consider is comparably large, but with six factors at work, it involves no more than 13 identified parameters (excluding $\Lambda_{Y,t}$).

The presence of V_t in expression (14) sets our approach apart from the USV settings, which explicitly prevent volatility factors from entering the cross section of yields. Collin-Dufresne, Goldstein, and Jones (2008) expose that such separation improves the ATSMs' fit to the spot volatility of yields. Considering our

empirical results, however, there appear to be few reasons—except statistical ones—for such constraint to hold in reality. In fact, there are at least two channels through which volatility variables could appear in the term structure, in particular at its long end. As stressed by Joslin (2007), one is the well-known convexity bias through which volatility is revealed in bond prices (see also Phoa, 1997). A second, and economically more important one, is the relation between the amount of uncertainty and the term premiums. In that the bond term premiums compensate for risks, they should be related to the changing amount of interest rate volatility.

In the benchmark model, we do not incorporate macroeconomic variables, and rely exclusively on the information contained in the yield curve itself. This choice is motivated by a simple fact. Namely, the Taylor rule—the basis of no-arbitrage macro finance—has been questioned as a model of the US monetary policymaking during the last two decades. And this for at least three reasons:[13] First and by definition, simple rules rely on a limited number of variables—maybe too limited to describe a complex economy like the one of the United States. Second, given the anticipatory nature of financial markets in general, and bond markets in particular, current or lagged values of a few macro variables may not provide sufficient guidance to the future economic conditions, let alone to the formation of the yield curve. Finally, the transmission of the monetary policy to the real sector in the US is strongly dependent on the capital markets. Since interest rates impact the value of equities, the stock market has important implications for the household wealth, consumer spending and investment. As such, it represents an important variable influencing the conduct of monetary policy. With model estimates at hand, in Section 6 we provide an empirical support to the economic roles of model-identified factors, and show that the information they aggregate transcends the content of the standard Taylor rule.

4. Model estimation

A practical implication of the weak link between yields and volatilities is that not all factors can be identified directly from yields. Thus, the backing-out technique—inverting Eq. (14) from observed yields to latent factors—would not work in our setting. We estimate the model on a weekly frequency ($\Delta t = \frac{1}{52}$) combining pseudo-maximum likelihood with a filtering technique. Since Y_t and V_t factors are unobservable, we express the model in a state-space form. At every date t, we filter the latent state by exploiting information both in yields and in volatilities.

4.1. Transition dynamics

The transition equation for Y_t is specified as an Euler approximation of the physical dynamics (6):[14]

[13] The speeches by Fed officials clearly demonstrate these points. See, for instance, the "John Taylor Rules" speech by the Fed's vice-chairman Donald Kohn on October 12, 2007.

[14] In Appendix B.2, we provide expressions for an exact discretization of the Y_t dynamics. We find that the use of the Euler scheme is virtually immaterial when $\Delta t = 1/52$, but offers a considerable increase in computational speed compared to the exact discretization. Therefore, the results presented here rely on the expression (16).

$$Y_{t+\Delta t} = \bar{\mu}_{Y,\Delta t} + \Phi_{Y,\Delta t} Y_t + u^Y_{t+\Delta t}, \tag{16}$$

where u^Y_t is a vector of heteroskedastic innovations $u^Y_t = \Sigma_Y(V_t)\sqrt{\Delta t}\epsilon_{t+\Delta t}$, and

$$\bar{\mu}_{Y,\Delta t} = \left(e^{\mathcal{K}_Y \Delta t} - I\right) \mathcal{K}_Y^{-1} \mu_Y \tag{17}$$

$$\Phi_{Y,\Delta t} = e^{\mathcal{K}_Y \Delta t}. \tag{18}$$

The transition equation for the matrix process V_t is obtained by an exact discretization of the dynamics (3):

$$V_{t+\Delta t} = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} + u^V_{t+\Delta t}, \tag{19}$$

where u^V_t represents a symmetric matrix of heteroskedastic innovations. Parameter matrices $\Phi_{V,\Delta t}$ and $\bar{\mu}_{V,\Delta t}$ are given as:[15]

$$\bar{\mu}_{V,\Delta t} = \int_0^{\Delta t} \Phi_{V,s} Q'Q \Phi'_{V,s} ds \tag{20}$$

$$\Phi_{V,\Delta t} = e^{M\Delta t}. \tag{21}$$

For ease of subsequent notation, we recast Eq. (19) in a vector form:

$$\text{vec}(V_{t+\Delta t}) = k\text{vec}(\bar{\mu}_{V,\Delta t}) + (\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\text{vec}(V_t) + \text{vec}(u^V_{t+\Delta t}). \tag{22}$$

Since the process V_t lives in the space of symmetric matrices, its lower triangular part preserves all information. Let us for convenience define two linear transformations of some symmetric matrix A: (i) an elimination matrix: $\mathcal{E}_n\text{vec}(A) = \text{vech}(A)$, where $\text{vech}(\cdot)$ denotes half-vectorization, (ii) a duplication matrix: $\mathcal{D}_n\text{vech}(A) = \text{vec}(A)$. Using half-vectorization, we define $\bar{V}_t := \text{vech}(V_t) = \mathcal{E}_n\text{vec}(V_t)$, which contains $\bar{n} = n(n+1)/2$ unique elements of V_t:

$$\bar{V}_{t+\Delta t} = k\mathcal{E}_n\text{vec}(\bar{\mu}_{V,\Delta t}) + \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n\bar{V}_t + \mathcal{E}_n\text{vec}(u^V_{t+\Delta t}). \tag{23}$$

Collecting all elements, we can redefine the state as: $S_t = (Y'_t, \bar{V}'_t)'$, whose transition is described by the conditional mean:

$$E_t\left(S_{t+\Delta t}\right) = \begin{pmatrix} \left(e^{\mathcal{K}_Y \Delta t} - I\right) \mathcal{K}_Y^{-1} \mu_Y + e^{\mathcal{K}_Y \Delta t} Y_t \\ k\mathcal{E}_n\text{vec}(\bar{\mu}_{V,\Delta t}) + \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n\bar{V}_t \end{pmatrix}, \tag{24}$$

and the conditional covariance of the form:

$$Cov_t\left(S_{t+\Delta t}\right) = \begin{pmatrix} Cov_t(Y_{t+\Delta t}) & 0_{n\times\bar{n}} \\ 0_{\bar{n}\times n} & Cov_t\left(\bar{V}_{t+\Delta t}\right) \end{pmatrix}. \tag{25}$$

[15] The closed form solution for the integral $\bar{\mu}_{V,\Delta t}$ is given by $\int_0^{\Delta t} \Phi_{V,s} Q'Q \Phi'_{V,s} ds = -\hat{C}_{12}(\Delta t)\hat{C}'_{11}(\Delta t)$, where

$$\begin{pmatrix} \hat{C}_{11}(\Delta t) & \hat{C}_{12}(\Delta t) \\ \hat{C}_{21}(\Delta t) & \hat{C}_{22}(\Delta t) \end{pmatrix} = \exp\left[\Delta t \begin{pmatrix} M & -Q'Q \\ 0 & -M' \end{pmatrix}\right].$$

See Van Loan (1978) for the proof.

The block diagonal structure in the last expression follows from our assumption that shocks in Y_t be independent of shocks in V_t. The respective blocks are given as:

$$Cov_t(Y_{t+\Delta t}) = \Sigma_Y(V_t)\Sigma_Y(V_t)'\Delta t \tag{26}$$

$$\begin{aligned} Cov_t\left(\bar{V}_{t+\Delta t}\right) &= \mathcal{E}_n Cov_t\left(V_{t+\Delta t}\right)\mathcal{E}_n' \\ &= \mathcal{E}_n\left(I_{n^2}+K_{n,n}\right)\left[\Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}'\otimes\bar{\mu}_{V,\Delta t}+k\left(\bar{\mu}_{V,\Delta t}\otimes\bar{\mu}_{V,\Delta t}\right)+\bar{\mu}_{V,\Delta t}\otimes\Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}'\right]\mathcal{E}_n', \end{aligned} \tag{27}$$

where $K_{n,n}$ denotes a commutation matrix (see e.g., Magnus and Neudecker, 1979). Buraschi, Cieslak, and Trojani (2008) provide the derivation of the last expression. Details are collected in Appendix B.

4.2. Measurements

We introduce two types of measurement equations based on yields (y_t^τ) and their quadratic covariation ($v_t^{\tau_i,\tau_j}$):

$$y_t^\tau = f(S_t;\Theta) + \sqrt{R_y}e_t^y \tag{28}$$

$$v_t^{\tau_i,\tau_j} = g(V_t;\Theta) + \sqrt{R_v}e_t^v. \tag{29}$$

Functions $f(S_t;\Theta)$ and $g(V_t;\Theta)$ denote model-implied expressions (14) and (15) corresponding to the observed measurements y_t^τ and $v_t^{\tau_i,\tau_j}$; Θ collects model parameters. $v_t^{\tau_i,\tau_j}$ is obtained from the high-frequency zero curve using estimator (1).

The weekly realized covariance matrix estimator is still quite noisy. It contains a significant portion of short-lived volatility which is not directly related to economic fundamentals but rather to various institutional effects, such as inventory management of primary dealers or calendar effects due to the Treasury bond auctions. To alleviate these effects, we construct every week a four week rolling realized covariance matrix. Such an adjustment makes it easier for the Kalman filter to distinguish between the noise and the fundamental volatility.

We assume additive, normally distributed measurement errors e_t^y and e_t^v with zero mean and a constant covariance matrix. In estimation, we use six yields with maturities of six months and two, three, five, seven and ten years. Yields share an identical standard deviation of measurement errors. Hence, their error covariance matrix is $R_y = \sigma_y^2 I_6$. Additionally, we include three volatility measurements which comprise variances of the two- and ten-year bond and the covariance between the five- and ten-year bond. The assumption of constant and identical measurement errors in yields is innocuous and intuitive. Lacking a similar prior for volatility measurements, we allow for different errors across equations, i.e. $R_v = \text{diag}(\sigma_v^i)$, where $i = 1,2,3$, and $\text{diag}(\cdot)$ denotes a diagonal matrix.

Implicit in our choice of measurements are two approximations. The first one is the assumption that the realized and the instantaneous covariance matrices of yields are equivalent. Indeed, neglecting e_t^v, Eq. (29) implies that: $\frac{1}{dt}\langle dy_t^\tau\rangle = v_t^\tau$ and $\frac{1}{dt}\langle dy_t^{\tau_i}, dy_t^{\tau_j}\rangle = v_t^{\tau_i,\tau_j}$, where $dt = 1/52$. In a strict sense and in absence of jumps, however, the estimator (1) converges to the integrated—rather than instantaneous—covariance

matrix of yields. We recognize that the measurement (29) is not exact, but on a weekly frequency the error due to the approximation can be assumed negligible.

Our assumption about the absence of jumps necessarily leads the second approximation. Even though extensions including discontinuities are readily possible, we specify the benchmark model as a pure diffusion. At the same time, the bi-power variation test applied to our high-frequency sample indicates that jumps are indeed present in the bond data. At the 99.8% confidence level, we find jumps on about 8% of days in our sample on average across maturities. These results are available upon request. Importantly, vast portion of jumps in the term structure appears at deterministic times (scheduled macro announcements). By nature, these jumps differ from the Poisson type of events. This motivates our choice of a purely diffusive model. Here, the weekly horizon comes in handy again: In the context of term structure models, one week appears a sufficiently long period to avoid the break-down of a pure diffusion.

4.3. Pseudo-maximum likelihood estimation

Despite linearity of the transition and measurement equations, the filtering approach introduced by Kalman (1960) is not directly applicable in our setting due to the non-Gaussian properties of the underlying state dynamics. In order to handle the non-Gaussianity we use the Unscented Kalman Filter (UKF), proposed by Julier and Uhlmann (1997), and recently applied in finance by e.g., Carr and Wu (2007) or Christoffersen, Jacobs, Karoui, and Mimouni (2009, CJKM). To approximate the conditional distribution, the UKF propagates the state through a set of deterministically chosen "sigma" points. Compared to the particle filter, it avoids costly simulations and thus offers a considerable gain in computational speed. This has a particular value for the estimation of multidimensional models like ours. Wan and van der Merwe (2001) argue that the approximation works well for purely diffusive settings. Indeed, CJKM (2009) show that the gain from using the UKF over the extended Kalman filter (EKF) is particularly visible for models with stochastic volatility. Details on the UKF implementation are provided in Appendix D.

Collecting all measurements in vector m_{t+1}, let $\hat{m}^-_{t+1}$ and $\hat{P}^-_{m,t+1}$ denote the time-t forecasts of the time-$(t+1)$ values of the measurement series and of their conditional covariance, respectively, as returned by the filter (for convenience 1 means one week). By normality of measurement errors, we can compute the quasi-log likelihood value for each time point in our sample:

$$l_{t+1}(\Theta) = -\frac{1}{2}\ln|P^-_{m,t+1}| - \frac{1}{2}\left(\hat{m}^-_{t+1} - m_{t+1}\right)'\left(P^-_{m,t+1}\right)^{-1}\left(\hat{m}^-_{t+1} - m_{t+1}\right), \tag{30}$$

and obtain parameter estimates by maximizing the criterion:

$$\hat{\Theta} := \arg\min_{\Theta} \mathcal{L}\left(\Theta, \{m_t\}_{t=1}^T\right) \quad \text{with} \quad \mathcal{L}\left(\Theta, \{m_t\}_{t=1}^T\right) = \sum_{t=0}^{T-1} l_{t+1}\left(\Theta\right). \tag{31}$$

with $T = 845$ weeks. The initial log-likelihood is evaluated at the unconditional moments of the state vector (see Appendix B for the expressions).

As a common problem in term structure modeling, the optimization of the loss function (31) is complicated by a high dimensionality of the parameter space, presence of multiple stochastic volatility factors, and complex interactions between parameters. Such circumstances leave little hope for standard local optimization methods, even when combined with a grid search. Therefore, to secure against local minima, we use the differential evolution (DE) algorithm designed for an efficient search of global optima in multidimensional, non-monotone, and multimodal problems (Price, Storn, and Lampinen, 2005). We confirm that the algorithm achieves the global minimum by repeating the estimation several times.

4.4. Identification

The latent nature of factors spells out the possibility that two distinct sets of parameters lead to observationally equivalent yields. To ensure econometric identification, we consider invariant model transformations of the type $\tilde{Y}_t = v + LY_t$ and $\tilde{V}_t = LV_tL'$, for a scalar v and an invertible matrix L. Such transformations result in the equivalence of the state variables, the short rate and thus yields (Dai and Singleton, 2000). If allowed, they can invalidate the results of an estimation.

To prevent the invariance, we adopt several normalizations for the physical dynamics of the process Y_t: (i) Setting $\mu_Y = 0$ allows to treat γ_0 as a free parameter. (ii) Restricting $\gamma_f = 1$ makes σ_f identified. (iii) Since both $\mathcal{K}_X$ and V_t determine interactions between the elements of X_t, they are not separately identifiable. We set $\mathcal{K}_X$ to a diagonal matrix, and allow correlations of the X_t factors to be generated solely by V_t. By the same token, the last row of matrix $\mathcal{K}_Y$, i.e. $(\mathcal{K}_{fX}, \mathcal{K}_f)$ is left unrestricted, as f_t does not interact with X_t via the diffusion term.

The identification of volatility factors V_t is ensured with three restrictions: (i) M is lower triangular and (ii) Q is diagonal with positive elements. (iii) The diagonal elements of Q are uniquely determined by setting $\gamma_X = \mathbf{1}_{n\times 1}$, where $\mathbf{1}_{n\times 1}$ is a vector of ones. These normalizations protect V_t against affine transformations and orthonormal rotations of Brownian motions. Finally, to guarantee the stationarity of the state, we require that the mean reversion matrices $\mathcal{K}_Y$ and M be negative definite. Due to the lower triangular structure of both, this is equivalent to restricting the diagonal elements of each matrix to be negative.

5. Estimation results

We discuss the parameter estimates and in-/out-of-sample model performance. Then, we study the dynamics of the filtered states and their respective contributions to the term structure of yields and volatilities.

5.1. Model performance

Table 3 provides parameter estimates for our model indicated by the label G_3SV_3 (meaning: three conditionally Gaussian plus three volatility factors). To obtain a comprehensive picture of its statistical properties, we consider two specifications: (i) a risk-neutral case, in which all parameters in $\Lambda_{Y,t}$ have been set to zero, and (ii) a risk premia case, with some of $\Lambda_{Y,t}$ parameters left free. On the level of explaining

yields, we compare the performance of the estimated models to their corresponding Gaussian three-factor counterparts (G_3SV_0), which have a proven track record on this front. We estimate the purely Gaussian models with the standard Kalman filter. This comparison puts to a test whether volatility factors, while beneficial for capturing second moments of yields, introduce a risk of misspecification on the side of yields.

That we consider models without risk compensation might surprise at first. Zero bond risk premiums are clearly untenable from an empirical perspective (e.g., Campbell and Shiller, 1991). Our motivation for this step is twofold. First, the ability of a model to describe the joint dynamics of yields and volatilities should be revealed from the structure of the underlying states rather than by adding new parameters. Indeed, we expect and demonstrate the risk premium specification to have a second-order effect on our model's performance in matching conditional volatilities. Second, while we certainly recognize that risk compensation exists in the bond market, we also find that the identification of price of risk parameters in $\Lambda_{Y,t}$ is a difficult task. Even for simple models, the estimates of premiums tend to have a low precision. With the inclusion of several parameters in $\Lambda_{Y,t}$, the search for a global optimum of (31) becomes slow and cumbersome. For this reason, we adopt a parsimonious form by judiciously selecting only two parameters in $\Lambda_{Y,t}$. We defer the details of this choice to Section 7.3. Interestingly, we find X_t's to be the only factors which have a significant impact on the term premiums.

Parameter estimates. After imposing identification restrictions, our preferred model presented in Table 3 has nine parameters which drive the yield curve (two of those describe market prices of risk), and six parameters which drive the term structure of volatilities. Even though the unobservability of states makes the interpretation of single estimates relatively uninteresting, several points stand out. Within each factor type at least one variable is more persistent and one faster-moving. While this observation may be common for the yield dynamics, our estimates indicate that it also holds for the volatility curve. Still, the most prominent difference between V_t states comes in their respective volatilities (vol of vol), as captured by the Q parameters. The volatility of the V_{11} and V_{12} state is just about a half of V_{22}. Such distinction is a first indication that a multivariate volatility is actually required by the data.

We also note large differences in the speed of mean reversion between Y_t and V_t factors: as expected, the autocorrelation coefficients decline much more rapidly for the volatility states. That variables generating yields have distinct statistical properties from those generating volatilities is consistent with the empirical findings of Section 2. Despite its intuitive appeal, this fact frustrates the volatility fit of many (especially three-factor) term structure models. But it is confirmed in our estimates of the V_t dynamics, specifically in the estimated degrees of freedom k. This single parameter controls the extent of non-Gaussianity present in the volatility factors: the higher the k the more Gaussian the dynamics.[16] Across all estimated versions, the model consistently selects $k = 2$. This low value reflects the need for factors with a pronounced right tail. Although such property suits the observed behavior of volatilities, it is less obvious how it could emerge in the term structure of interest rates. We interpret it as supportive of our design that separates distributional qualities of yields from those of volatilities.

[16] Buraschi, Cieslak, and Trojani (2008) discuss the role the degrees of freedom parameter in the V_t dynamics. See also Appendix B.1 for the properties and construction of the V_t process. The representation of V_t as the sum of k outer products of the multivariate Ornstein-Uhlenbeck process provides the key intuition for the role of k.

In-sample fit. The last section of Table 3 provides the log-likelihood values for each estimated specification. For comparison with the purely Gaussian benchmark, we split the log-likelihood implied by our model into two portions measuring the fit to yields and to volatilities, respectively. A look at the numbers indicates that all versions have similar abilities in explaining yields. In particular, the improvement in volatility modeling offered by G_3SV_3 spells no sacrifice in terms of fitting interest rates (see Fig. 7). Across a spectrum of maturities, the model replicates the observed weekly dynamics of yields with a high accuracy. Simultaneously, it is able to explain well the evolution of the second moments. Fig. 8 superimposes the observed and model-implied behavior of the level and the slope of the volatility curve, and the covariance between the five- and ten-year yield. Level and slope are proxied by the two-year volatility and the spread between the ten- and two-year volatility, respectively. The plot shows that the model-implied dynamics track observed quantities very closely in terms of magnitudes, persistence, and signs. By being consistent with both the level and the slope, our setting provides a good description of the volatility at the short and the long end of the curve.

[Fig. 7 and 8 about here.]

Table 4 confirms these conclusions. We provide two yardsticks of in-sample model performance: the root mean squared errors (RMSE) and the percentage of variation in yields and volatilities explained by the model. The latter is measured with the R^2 coefficient from a regression of the observed on the model-implied dynamics. The RMSEs show that on average the model misses the true yield by about two basis points. As such, it is able to explain virtually the total of observed variation in yields. Its performance is very similar to the purely Gaussian case. The match to the volatility dynamics is less perfect than that to yields, still we are able to explain from 94% to 98% of variation in the second moments dynamics.

Yield forecasting performance. When assessed in sample, our model can conceal distortions coming with a richer form of the state space, if misspecified. This should be revealed in a quick deterioration of its performance out of sample. To check this possibility, we re-estimate the model on the data from January 1992 through December 2004, and use the remaining part of the sample (January 2005 through December 2007) for out-of-sample evaluation. For comparison, we take the same approach for the purely Gaussian setting. With new estimates, we obtain predictions of yields over horizons of one, four, 12 and 52 weeks. Table 5, panel *b* summarizes the results. The yield forecasting ability implied by our model is very close to the purely Gaussian case. It is well-known, yet unfortunate, that the existing term structure models fail in outperforming the random walk predictions out-of-sample (see e.g. Duffee, 2002; Moench, 2008). While at the very short forecast horizons (i.e. one week) both settings are outpaced by a simple random walk, they do a decent job in forecasting yields over longer term (e.g. one year). The upshot is that the additional volatility factors do not induce instability in the model performance.

Volatility benchmarks. In terms of explaining the volatility curve, the model we propose has no direct benchmark in the literature. The interpretation of the existing results is intricate as approaches differ by the data and sample period used (Treasury or swap rates, with or without interest rate derivatives), restrictions imposed (with or without USV), and estimation methods applied. These discrepancies often lead to opposing

conclusions, thus making a direct comparison difficult.[17] We take $A_1(4)$ models estimated by CDGJ (2008), and Thompson (2008) as a benchmark. A consensus emerging from those papers is that the $A_1(4)$ model can do reasonably well in matching volatility at the short end of the curve, but its performance deteriorates quickly with the yield's maturity. For instance, CDGJ report that instantaneous volatility of the ten-year yield implied by the $A_1(4)$ model is 18 times less volatile than the rolling window volatility estimates, suggesting that more than one state variable is needed to capture the term structure of volatilities. The corresponding number returned by our model is 7%, and is coupled with the ability to explain 98% of the variation in the slope of the volatility curve.

Similar to the yield forecasting exercise, we compare volatility forecasts to a random walk. Given the fast decay of volatilities, their forecasts at horizons longer than a couple of months are not meaningful. Table 5, panel *b* summarizes the results. Our model performs particularly well for four and 12 week horizons, even though the out-of-sample period (2005:01–2007:12) privileges the random walk due a consistently low and stable level of volatilities. Across the maturity spectrum, the volatility of the five- and ten-year yield is predicted with the highest accuracy. This outperformance is explained by the presence of an additional volatility state which, as we argue below, is dedicated to capturing precisely this segment of the curve.

5.2. Filtered states

In that the model gives a reliable description of the data, it should reflect the key drivers underlying the formation of the curve. This section discusses their properties and roles.

Fig. 9 plots the state variables extracted by the unscented Kalman filter. Panels on the left display the evolution of the three yield curve factors, $Y = (X_1, X_2, f)$, those on the right show the covariance factors $\bar{V} = (V_{11}, V_{12}, V_{22})$.

[Fig. 9 about here.]

This autonomous behavior of the two factor groups reveals no obvious candidate variable that could be acting equally on yields and on their second moments. This does not mean, however, that factors lack common interpretation. Quite the opposite: when fed with the data, the model puts the right pieces of its structure into the right places. Most importantly, the roles it assigns to the X_t states correspond neatly with the covariance interpretation of the respective volatility factors V_t.

To see this, it is useful to study the effects that different variables have on the yield curve. Fig. 10 plots the responses of yields to shocks in each element of the state vector. Panels on the left $(a1, a2, a3)$ display the reaction of the curve to perturbations of the Y_t variables; panels on the right $(b1, b2, b3)$—to shocks in V_t variables. In each subplot, the solid line depicts the yield curve when all variables are held at their

[17] The best example comes from comparing the results reported by Thompson (2008) versus those reported by CDGJ (2008). Both papers estimate the $A_1(4)$ model with and without USV using swap and LIBOR rates, yet reach different conclusions. Thompson (2008) shows the model can track volatility dynamics similarly well whether or not the USV is imposed. CDGJ demonstrate, in contrast, that the USV restriction significantly improves the model's performance. By relying on the spot and option price data, the results for the $A_1(4)$ and $A_2(4)$ specification reported in Joslin (2007) are less well suited for drawing comparisons here.

unconditional means. Instead, the dashed and dotted lines plot the yield curve response when a given factor is set to its 10th and 90th percentile, respectively. Let us for the moment focus on the impact of the yield curve factors. Several points emerge. The effect of the f_t state is most pronounced at the short end of the curve. A downward (upward) shift in f_t moves the short yield significantly below (above) its unconditional mean. The effect diminishes with the maturity, and is consistent with the filtered time-series pattern of f_t which closely tracks the shortest yield (see Fig. 9*c*). The two X_t factors influence the longer segment of the curve. X_2 is most active at intermediate maturities (two–three years), inducing changes in the curvature. X_1 acts predominantly at long maturities, thus changing the slope.[18] The identification of X_t states as those impacting on longer maturities concurs with the form of the market price of risk we adopt (see discussion in Section 7.3). Recall that our preferred specification chooses the X_t factors to be informative about risk compensations, which are most pronounced in the longer segment of the curve.

The roles of X_t factors are in line with the identified volatility states. Indeed, V_{11} generates the volatility at longer maturities, while V_{22} is responsible for capturing the shorter end of the volatility curve. To support this conclusion, Table 6 reports t-statistics obtained by regressing observed yield volatilities and covariances on the elements of the state vector. The significance of V_{11} and V_{22} factors for explaining the observed volatility curve has an opposite pattern: The effect of V_{22} is decreasing with the maturity, in contrast the effect V_{11} is increasing with the maturity. The out-of-diagonal element V_{12} is overwhelmingly significant, but has the largest impact on the yield covariance dynamics.

5.3. Short and long-end volatility: two episodes

The split of volatility factors into the short- and long-end components is also evident in their filtered dynamics. The long-term factor V_{11} reflects major moves in the volatility of the ten-year yield. The short-term factor V_{22}, in turn, captures the two-year volatility. To illustrate this point, let us consider two salient moments in the volatility history (marked with circles in Fig. 9). In the last months of 2003, the large move of the V_{11} factor mirrors a rapid increase in volatility that affected the long end of the curve, but was barely visible at its short end (panel *d*). The main trigger for the elevated volatility in this period came from the real part of the economy. In June, the Fed lowered rates to 1% and was communicating no tightening for a foreseeable future. This anchored certainty about the short end of the curve. At the same time, the second half of 2003 welcomed first signs of recovery, but they continued to be mixed: Increases in retail sales and productivity gains mingled with still weak payroll figures. Investors were struggling to interpret the implications of these numbers for the long term, spurring a new thrust of volatility at the long segment of the curve.

Quite a different, but consistent, story is related to an earlier volatility episode in June 1995 highlighted in panel *f* of Fig. 9. This moment, which marks the largest move in volatility of the two-year rate, is manifestly captured by the V_{22} dynamics. Yet, it does not appear more than a usual blip either in the other two factors, or in the volatility of longer yields. Accordingly, this time the trigger to the event had a short-end nature.

[18] Since factors are latent, we do not discuss the direction of their impacts when at the tenth or 90th percentile. The particular sign is a consequence of the identifying normalization we impose. f_t factor is an exception, given its close resemblance to the short rate.

In mid-1995 the monetary policy easing was imminent. The market was expecting that the Fed would take an aggressive pace to prop up the economy, as indicated by the BlueChip Economic Indicators survey. In May the expectations of the three-month T-bill rate plunged remarkably by over 1%. Though the Fed did ease in the end, it did so by much less then expected. This raised confusion at the short end of the curve, but left the long-end almost intact.

However incidental, the two volatility episodes expose well the roles of factors that move the volatility curve. In particular, they back our finding of long- versus short-run volatility components. To provide further support for this interpretation, in Section 6 we establish more rigorously the link between factors and observable economic quantities. Before doing so, however, we turn to exploring whether and how the volatility states influence the yield curve.

5.4. Are volatility factors revealed by the yield curve?

To understand how our model answers this question, let us focus again on Fig. 10. Its right-hand panels $(b1, b2, b3)$ plot the response of the yield curve to large perturbations in the volatility states. Relative to the strong reaction induced by the Y_t factors, the impact of the volatility states on the cross section of yields turns out much weaker. Even spectacular shifts in the V_{22} and V_{12} factors—recall that we consider their tenth and 90th percentile values—do not exert a visible effect on interest rates of any maturity. The impact of V_{11} is only revealed at maturities beyond five years, still it does not exceed a few (on average seven) basis points. This fact has an intuitive meaning: Abstracting from the potential volatility impact on the term premiums, we would expect the long-term volatility factor V_{11} to show up at longer maturities via the convexity effect. Since this effect is mainly present in long yields, the response to V_{22} should be less pronounced, or even negligible.

To illustrate this effect over time, Fig. 11 plots the model-implied yields $y_t^\tau = -\frac{1}{\tau}\{A(\tau) + B(\tau)'X_t + Tr[C(\tau)V_t]\}$ against hypothetical yields produced by removing V_t, i.e. $y_t^\tau = -\frac{1}{\tau}\{A(\tau) + B(\tau)'X_t\}$. The graph makes our earlier point more explicit: Volatility factors contribute a tiny margin to the dynamics of yields. The curve is virtually indistinguishable whether or not V_t factors are included. With $\tau \to 0$, the impact of V_t is excluded by construction of the short rate in Eq. (4). It turns out, however, that also in matching longer yields the model chooses not to employ the volatility states.

Our estimates suggest that matrix of factors loadings, whose rows are given by $\left[-\frac{B(\tau_i)}{\tau_i}', -vech\left(\frac{C(\tau_i)}{\tau_i}\right)'\right]$, $i = 1, \ldots, 6$, is close to singular (condition number of 9000). It is thus hardly possible to back out a complete set of volatility states from yields. Still, we might be able to recover at least one of them. We reduce the dimension of the loadings' matrix to four and select only factor V_{11}, i.e. $\left[-\frac{B(\tau_i)}{\tau_i}', -\frac{C_{11}(\tau_i)}{\tau_i}\right]$. This matrix becomes easily invertible (condition number of 16), but it is still difficult to empirically disentangle the elements of X_t from the volatility factor as both have to be identified from the same long-end yields. The quality of inversion depends on the subset of yields chosen, which explains why the backing-out technique has troubles identifying volatility states.

5.5. Filtered factors versus principal components

Finally, we would like to understand how our extracted factors relate to the usual yield curve metrics—the orthogonal principal components. How does their interdependence and the presence of V_t states impact the interpretation? Table 7 presents contemporaneous regressions of filtered states on the five principal components (PCs). Each of our yield curve factors is strongly related to two PCs. Factor X_1 is linked to PC_1 (level) and PC_2 (slope) exactly as factor f is. But while X_1 loads with the same sign on the level and slope, factor f loads positively on level but negatively on the slope. Our intuition is that a positive shock to f—a positive monetary policy shock—is linked to a higher level but a lower slope due to, e.g., the decreased inflation expectations. Alternatively, one can interpret factors X_1 and f as two levels, a long- and a short-term, respectively. These levels respond to economically distinct shocks (see Section 6) and are connected in a nontrivial way through the monetary policy transmission mechanism. Interestingly, support to our two-level interpretation of factors is given in a recent Wall Street Journal article by the ex-Fed Chairman Alan Greenspan, as he discusses the decoupling of the monetary policy from the long-term rates after the year 2000 (Greenspan, 2009). Factor X_2 is related to the level and to PC_3 (curvature) with the opposite signs, and captures the monetary transmission up to the medium term.

Our volatility states pick up tiny movements in the curve embedded in principal components of higher order (PC_4 and PC_5). But not only. V_t's are also related to the slope (PC_2). Empirical evidence in Section 2 helps explain this fact: Monetary easing in our sample, raising the slope of the term structure, coincides also with a pronounced boost of the volatility curve (see e.g., Fig. 1). Therefore, and despite their very different roles for the yield curve, both volatility states and the slope happen to respond to a common factor. In terms of higher-order PCs, the long-end volatility V_{11} and the covariance term V_{12} load with large and significant coefficient on the PC_4. While more weakly linked to the filtered states, PC_5 is also significant across the board.

This section leads us to several conclusions. There appears to be little evidence that one and the same variable generates both the term structure of interest rates and of interest rate volatilities. In combination with the knowledge that three factors are needed to explain just yields, such discrepancy is a tall order for any three-factor model. Precisely for this reason, the USV restriction may come in handy as a tool for separating roles of factors in low-dimensional settings. Without such restriction, three-factor models face misspecification risk on the volatility modeling front. Consequently, more factors are needed, and just one more may not be enough. There is little support that volatility factors can be identified from the cross section of yields. Rather, help of extra information from interest rate options, instantaneous volatility proxies (like the ones we apply), and estimation methods (filtering) is required.

6. Factors and surveys

The current section takes two steps to investigate the economic content of factors. First, we seek additional evidence for our interpretation of latent states using forecasts about the future direction of US interest rates. Then, we relate the filtered dynamics to expectations and uncertainties about macro aggregates. We rely

on surveys rather than realized numbers motivated by the notion that prices and state variables shall reflect economic prospects better than past events. Our monthly survey data are spliced from two sources. To obtain forecasts of macro variables, we use the BlueChip Economic Indicators (BCEI) survey. BlueChip Financial Forecasts (BCFF), in turn, supply us with predictions of interest rates. As evident in the transcripts of the FOMC meetings, these surveys are regularly used by policymakers at the Fed to read market expectations. Appendix A.4 provides details about the survey data including their timing within month. We use responses of individual panelists to construct proxies for the consensus forecast and for the uncertainty. Each month, the consensus is computed as the median survey reply. The uncertainty is measured with the mean absolute deviation of individual forecasts.

6.1. Expectations and uncertainties about the yield curve ...

Our previous findings suggest an important distinction between factors. To prop this conclusion without resorting to price data used in our estimation, surveys about the evolution of US interest rates are useful. We construct monthly observations of filtered states by averaging weekly series within each month, and merge them with the most contemporaneous survey results. Table 8 reports regressions of factors on expectations and uncertainty about US interest rates, respectively denoted with E and σ. Every coefficient is additionally equipped with a qualifier $(hQ1, hQ4)$ indicating the horizon for which the forecast of a given yield was made (next quarter versus four quarters ahead). The different horizons turn out to be mutually exclusive, i.e. we do not observe situations, in which both $hQ1$ and $hQ4$ are jointly significant. In Table 8, we preserve only significant explanatory variables.

Let us consider the regressions of yield curve factors (X_1, X_2, f) first. Consensus rate forecasts are able to explain 76%, 78% and 93% of their variation, respectively. The significance of respective loadings confirms the split of yield curve factors into long-, intermediate- and short-run. Accordingly, X_1 is most strongly related to the forecasts of the ten- and 30-year rates, X_2 captures the two-year yield forecast, and f is almost entirely linked to the consensus federal funds rate. This consistency carries forward onto the volatility states. The long-run volatility factor V_{11} is associated with the uncertainty surrounding the long end of the yield curve from above five to 30 years. In contrast, the short-run volatility factor V_{22} loads only on the uncertainty about the two-year yield.

Interestingly, the short- versus long-run effects have an additional dimension revealed in the pattern of the forecast horizon identifiers. We find that X_1 and its volatility V_{11} are both most strongly linked to yield predictions at longer horizons $(hQ4)$. This reverses for factors which we diagnosed as shorter-run. Table 8 shows that shorter-lived predictions $(hQ1)$ play a dominant role for explaining all state variables except X_1 and V_{11}. Incidentally, distinct segments of yield and volatility curves appear to capture how the market assesses economic prospects and uncertainties at different horizons into the future.

6.2. ... and the macroeconomy

While it seems natural that different regions of the yield curve are driven by different underlying forces, current discussion suggests that the same holds true for their volatilities. Therefore, it is tempting to

understand which economic variables might stand behind those moves. Table 9 presents the regressions of factors on consensus and uncertainty proxies about macroeconomic quantities. For clarity, we include only significant loadings. The macro variables we use represent three standard domains: (*i*) real activity is captured by real GDP growth (RGDP), industrial production (IP), unemployment (UNEMPL), and housing starts (HOUST); (*ii*) the federal funds rate (FFR) describes the stance of the monetary policy, and (*iii*) inflation is reflected through the CPI forecasts. The combination of expectations and uncertainty proxies constructed from these measures turns out to be overwhelmingly informative about our filtered states. We are able to explain up to about 90% and 50% of variation in the yield curve and volatility factors, respectively.

The picture emerging from those regressions has many angles, but some general observations can be made. The distribution of loadings among the yield curve factors confirms their distinct responses to the economic environment. The long-term yield factor X_1 is strongly driven by the expectations about the real sector, with unemployment emerging as its key mover. Being an intermediate state, X_2 straddles both the real and nominal effects and puts additional weight onto uncertainties. X_1 and X_2 load both on consensus and uncertainty about the real GDP. Still, the magnitude of their reactions to either component is different, and stronger for X_1. Finally, in addition to being almost completely explained by the monetary policy, the short-end factor f_t responds somewhat to the CPI inflation and the current stance of the real activity. The latter is timely captured by the industrial production. Not surprisingly, the monetary policy is significant for all yield curve states, but its impact diminishes with the state's maturity, telling us how the monetary impulse is transmitted to the longer segments of the curve.

Macroeconomic conditions are also reflected in the V_t states. V_{11} shows the strongest response to measures of real activity, and is pronouncedly influenced by the real GDP growth, the uncertainty about unemployment, and expectations on housing starts. Such configuration is consistent with its interpretation as the long-run volatility state. Quite different variables have explanatory content for the short-term volatility state, instead. V_{22} is linked to the uncertainty about inflation, industrial production and monetary policy. Notably, shocks to these variables are almost contemporaneously observed and rather short-lived, thus providing a contrast to more sticky metrics as the real GDP.[19] Indeed, variables inducing long- versus short-run volatility movements form virtually disjoint sets. Expectations of a monetary policy easing, for instance, spur an increase in V_{22}, but do not appear to have any visible effect on V_{11}. This gives a meaning to our finding in Section 2 that volatilities of shorter yields respond more to interest rate cuts than do volatilities of longer yields: It is the short-run volatility component that provides for the asymmetry.

To better understand the difference between the volatility states, in panels *a* through *e* of Figure 12 we plot their impulse-responses to various sources of macro uncertainty. The respective reactions differ in terms of magnitude and persistence as well as types of variables which are important. Uncertainties about FFR and inflation have an effect on short volatility which can last up to six months, but their impact on the long volatility remains negligible. This pattern reverts for the real GDP uncertainty, whose role persists at

[19] Ranking inflation as short-lived may be surprising. This can be partially explained by the sample period we use (1992:01–2007:12). Looking at the term structure of inflation expectations provided by BCFF surveys, we notice that while short-run forecasts (for the current quarter) are surprisingly volatile, their volatility dies out very rapidly. Starting already from one-quarter-ahead, inflation expectations become extremely smooth. In his speech on January 3, 2004 at the American Economic Association meetings, Mr. Bernanke mentions this feature explicitly.

the long end but is only contemporaneously important at the short end. Interestingly, V_{11} has a significant impact on V_{22}, but not vice versa, as visible in panel f. The intuition for this result comes with a simple example. Imagine a situation in which uncertainty about the GDP growth picks up, and elevates V_{11}. The same news likely spurs uncertainty about near-term activity measures such as industrial production that affect V_{22} on the way. However, inverting the scenario and recognizing a higher duration of the long-term volatility, we do not necessarily expect a symmetric effect in the opposite direction to take place.

Admittedly, interactions between model-implied factors and macro variables deserve a deeper discussion than the one we venture here. A simple analysis is unlikely to reflect the complexity of lead-lag relations between different variables. While exploring these relations in depth is beyond the current scope of this paper, here we provide one example. Housing starts are significant across the board, and as a sole variable simultaneously appear in the short- and long-run volatility components. The key to understand this fact lies in the role that housing plays for the real economy and thus monetary policy in the US. The slowdown in the housing market is thought to have a significant effect on the real GDP growth via reductions in consumer spending (negative household wealth effect and decline in mortgage equity withdrawals, see e.g. Mishkin (2007)). In conditions of contained inflation as in our sample, the Fed is expected to counteract this negative development by lowering the interest rates. This double-edged sword effect of housing on the real economy and monetary policy shows that the apparently confusing presence of this variable in both V_{11} and V_{22} can actually be consistent with the overall interpretation of these factors.

Turning to the interpretation of the out-of-diagonal covariance factor, we find that—under our identification scheme[20]—V_{12} is highly positively linked to the realized correlation of the long-to-medium part of the yield curve (see Figure 2, panel c). Contingent on this observation, we can interpret its signs in survey regressions. V_{12} is associated with uncertainty measures about the real GDP, unemployment, housing starts, and monetary policy, all of which load with a negative sign. An increase in uncertainty about the macroeconomy implies a decline in V_{12}, and a lower correlation between the long and medium part of the yield curve. Intuitively, the different segments tend to move more independently when the curve changes shape. This, in turn, usually occurs in times of increased uncertainty such as April–June 2001, which is the only NBER-proclaimed recession in our sample (see our collection of major curve moves in Table 1). V_{12} captures precisely those moves.

7. Discussion and robustness

This section discusses the robustness and efficiency of the realized second moment estimator (1) that are critical for our results. We also address the market price of risk specification, and its selection approach we follow in estimation.

[20] Our latent states are subject to identification restrictions which preclude direct interpretation of factor signs (except for V_{11} and V_{22}, which are always positive). Recognizing this fact, we seek to establish factor signs that give their natural interpretation in terms of observable quantities.

7.1. Realized covariance matrix estimation

In case of asynchronous trading, the realized covariance matrix estimator defined in Eq. (1) can be biased toward zero (see e.g. Hayashi and Yoshida, 2005; Audrino and Corsi, 2007). The bias is to a large extent generated by the interpolation of non-synchronously traded assets, and its severity depends on the difference in liquidity of the assets considered.[21] Hayashi and Yoshida (2005, HY) propose a covariance estimator which corrects for the bias in (1). The estimator sums up all cross-products of returns which have an overlap in their time spans, and thus no data is thrown away. The covariance of two bond yields reads:

$$RCov_{i,j}^{HY}(t, t+h; N_i, N_j) = \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \left(y_{t_i + \frac{kh}{N_i}} - y_{t_i + \frac{(k-1)h}{N_i}} \right) \left(y_{t_j + \frac{lh}{N_j}} - y_{t_j + \frac{(l-1)h}{N_j}} \right) \mathbb{I}\left(\tau_i \cap \tau_j \neq 0\right) \tag{32}$$

where τ_i and τ_j denote the interval of the return on the first and second bond, respectively.

To verify the robustness of our realized covariance estimator (1), we implement the HY approach for the realized covariance of ten- and five-year bond. Both estimators deliver very similar results in terms of magnitude and covariance dynamics. They are highly correlated (90%) and the t-test for the difference in means does not reject the null that $\mu_{\text{outer}} = \mu_{HY}$ (p-val $= 0.57$).

There at least two reasons why we stick to the simple outer-product realized covariance estimator (1). For one, estimators of Hayashi and Yoshida (2005), Audrino and Corsi (2007) are not directly applicable in our case because for the construction of the zero curve we require a synchronized set of yield changes. More importantly, on-the-run Treasury bonds are largely homogenous in terms of liquidity, which is well proxied by the average number of quotes/trades per day reported in Table A-1 of Appendix A.2.

7.2. Alternatives to the realized covolatility estimator

Instead of the realized second moments, one could resort to other volatility measures. To assess the gain from using high-frequency data, we consider two alternatives: (*i*) the DCC-GARCH model proposed by Engle and Sheppard (2001), and (*ii*) the realized covolatility computed from the daily data.

In a first step, we estimate the DCC-GARCH(1,1) specification on demeaned yield changes at daily, weekly and monthly frequency. Estimation results are available upon request. We see several reasons to prefer the realized estimator to the GARCH. Most of all, the realized covolatility has a richer economic content that is embedded in the intraday yield movements. The residuals from a regression of realized volatility on the GARCH volatility still carry significant economic information. Almost 30% of the variation in residuals can be explained by macroeconomic surveys. Second, the realized measure is nonparametric and thus not prone to a misspecification. Third, GARCH estimates cannot be easily aggregated to lower frequencies (Drost and Nijman, 1993). Based on monthly data, GARCH estimates are imprecise compared to the realized volatility. Thus, if using GARCH, our analysis in Section 6 would be less reliable. Finally, the DCC-GARCH model has difficulties in matching the abrupt changes in correlation evident in the realized correlation estimates.

[21] Audrino and Corsi (2008) offer a thorough discussion of the bias in realized covariance.

In a second step, we compare our realized covolatility measure to the one estimated on the daily data. This exercise gives us a view on the additional value coming from the intraday data. The daily estimator can explain a significant part (around 70%) of the variation in the realized covolatility. However, similar to above, the residuals from the regression of realized volatility measures on the volatility from the daily data are still informative. Up to 20% of their variation can be explained by the economic forecasts. The importance of including intraday observations increases with maturity.

Those comparisons show that the reason for using high-frequency data is beyond the standard statistical efficiency arguments. The incremental economic content of intraday observations is consistent with our interpretation of realized volatility as providing complementary piece of information towards yields themselves.

7.3. Specification of bond premiums

Both economically and statistically, it is hard to identify all parameters of market prices of risk Λ_Y driving excess returns of bonds. Therefore, we reduce the number of parameters to be estimated to a necessary minimum. This section describes our selection procedure.

We write log excess holding period return from a simple strategy of buying an n-year bond at time t and selling it in one year, i.e. for weekly data at $t+52$, the strategy is financed by rolling over one year bond:

$$rx_{t+52}^{(n)} \equiv p_{t+52}^{(n-1)} - p_t^{(n)} - y_t^{(1)}, \tag{33}$$

where $p_t^{(n)}$ denotes log of bond price $P_t^{(n)}$, and n is expressed in years. To understand how excess returns are related to the yield curve, we regress the realized excess returns of two- to ten-year zero bonds defined in Eq. (33) on the vector of filtered states $(Y_t', vech(V_t)')'$:

$$rx_{t+52}^{(n)} = \alpha^{(n)} + \beta_Y^{(n)} Y_t + \beta_V^{(n)} vech(V_t) + \epsilon_{t+52}^{(n)}. \tag{34}$$

Because at this point we are agnostic about the structure of model-implied risk prices Λ_Y, we obtain factors for the above regression from a model estimated with zero market prices of risk for all shocks. There is a reason for keeping all states in the predictive regression: The six factors are in the information set of an investor at time t, i.e. having high-frequency bond data the investor can filter all of them. While those regressions explain between 41% (two-year) and 33% (ten-year) of variation in realized excess returns, based on Hansen-Hodrick standard errors only X_1 and X_2 turn out statistically significant at the 1% level.

The part of bond excess return that can be predicted using the yield curve information has the following form:

$$E_t\left[rx_{t+52}^{(n)}\right] = \alpha^{(n)} + \beta_Y^{(n)} Y_t + \beta_V^{(n)} vech(V_t). \tag{35}$$

In a first step, we determine the main sources of the time-variation in excess returns. Similar to Cochrane and Piazzesi (2008), we stack expected excess returns across maturities and extract principal components. A single factor—the first PC—can explain about 98% of the return variation and its loading increases with the bond maturity. Importantly, around 80% of the movements in this dominant PC can be captured by X_1 and

X_2. This finding has immediate implications for modeling premiums. In particular, the return-forecasting factor is largely spanned by our first two yield curve states.[22] Hence, we let X_1 and X_2 drive the variation in excess returns, both of which exhibit stochastic volatility and a nontrivial correlation.

In a second step, to establish which shocks are priced, we estimate a VAR system for the filtered yield curve factors Y_t, and compute covariances of factor shocks with the realized excess returns across maturities. Covariances of shocks to X_1 and X_2 with the realized excess returns are much larger than those of f, and in both cases rise with maturity: covariance of X_1 increases linearly and of X_2 has a hump-shape. This pattern suggests that shocks to factors X_1 and X_2 are priced. Therefore, from Eq. (7), we specify:

$$\Sigma_{Y,t}\Lambda_{Y,t} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda^1_{11} & 0 & 0 \\ 0 & \lambda^1_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ f_t \end{pmatrix}. \tag{36}$$

In order to corroborate this choice, we turn again to regressions of yield factors on surveys (Section 6.2, Table 9). Yield factors are to a large degree explained by the expectations and uncertainties about the macro conditions. As such, their innovations should subsume shocks in those expectations and uncertainties. Factor f captures rather short-lived shocks stemming from industrial production and inflation. Instead, the likely candidates for impacting premiums on longer-duration bonds in our sample—output growth, unemployment and housing market—enter the yield curve precisely through factors X_1 and X_2.

While with specification (36) we aim to provide a decent description of the time-varying portion of premiums, we remain moderately optimistic about its ability to capture the full economic content of expected excess returns. Several symptoms foster our caution. For instance, the residuals from regressions (34) of realized excess returns on yield curve factors (or similarly on the Cochrane-Piazzesi factor) are highly autocorrelated (0.96). This points to the existence of other return predictors outside the yield and volatility curve. Moreover, we find it possible to significantly increase the R^2's of those regressions by including the cross-products of yield curve factors. Unfortunately, such market prices of risk are hard to implement in our model without sacrificing its tractability. The convenient affine property of both the physical and risk neutral dynamics is potentially traded off for some misspecification on the side of term premiums. With our approach to selecting Λ_Y, we try to introduce a minimal misspecification while following the basic economic intuition, and ensuring stable parameter estimates. Indeed, including the unconstrained matrix Λ_Y can considerably hamper the optimization of the loss function (31), and thus distort the identification of latent states.

8. Conclusions

We study and model the joint behavior of the yield and the volatility curve. Understanding the interest rate volatility requires an additional effort as its dynamics is hidden from the perspective of the cross-section

[22] Recall that the states we filter out differ from the traditional PCs extracted from yields. Therefore, the explanatory power of X_t's for the return-forecasting factor is not in contradiction with the results of Cochrane and Piazzesi (2005, 2008), who show that their factor goes beyond the standard PCs.

of yields. However, with 16 years' worth of tick-by-tick Treasury bond transaction data, we obtain a good and contemporaneous description of the evolution of yields and their realized covariances. Several facts emerge from the analysis of the realized series: The level, slope and curvature portfolios comove in a nontrivial way, so that an unconditional PCA neglects potentially important interactions between states forming the curve. Interest rate volatility curve does not move on a single factor, rather several states are needed to comprise its dynamics. Finally, while the yield-volatility link is weak on average, we also show that it is time-varying and subject to the influence of the monetary policy.

We take the empirical observation as a clue for the model design. To decompose the two objects, we develop a no-arbitrage model able to capture jointly the observed dynamics of yields and their volatilities. While additional restrictions (e.g. the USV) are not called for by the data, the model's empirical performance relies on two elements: one is the multivariate volatility structure separating factors in volatilities from those spanning yields, the latter is the identification of volatility states with support of the filtering and of the realized covolatility proxies.

The model disentangles clear long- and short-run volatility components which correspond to the respective yield curve states, and consequently, yield durations. An additional covolatility factor accommodates the interactions between the long and intermediate region of the curve. Neither of those states manifests itself in the cross section of yields: If at all, the most pronounced impact comes from the long-end volatility, but it does not exceed a few basis points and is only visible at maturities of beyond five years.

We start by being agnostic about factor names, and only after having shown their explanatory content, we link them to the forward-looking economic quantities using surveys. Factors differ by shocks to which they respond. The short-run volatility reflects the impacts of changing expectations and uncertainties about short-lived, at least in our sample, economic variables (e.g. inflation, monetary policy, and industrial production). The longer-run volatility is associated with more sticky measures of macroeconomic conditions (e.g. real GDP growth, unemployment). The covolatility term is sensitive to uncertainties about the real activity and the Fed, all of which contribute negatively to the correlation between the middle-to-long segment of the yield curve.

As such, the latent factors do not appear to be void. Each of them compounds influences of several variables, and shows how they are transmitted throughout the term structure. Restricting the discussion up front to few macro variables, instead, could obscure some important effects uncovered by the current approach. Still, while yield curve states are well explained with the combination of surveys, a large portion (about half) of volatility remains nonfundamental from the perspective of the economic variables we consider. Interactions with the stock and other markets, that are beyond the scope of this study, are potential ingredients contributing to the unexplained part.

This work can be extended in several directions. Most importantly, we could spend more time on specifying and studying prices of risk, and their relationship with the amount of risk in interest rates. Though this aspect is of second-order importance when explaining the yield volatility curve, it is key for understanding the bond term premiums. This direction is the focus of our research in future.

Appendix A.

This Appendix gives a brief description of the high-frequency Treasury data, our zero curve construction methodology, and macroeconomic surveys.

A.1. GovPX and BrokerTec

Table A-1 reports some basic statistics on the Treasury bond transaction data in the period 1992:01 through 2007:12. For the GovPX period (1992:01–2000:12), we report the average number of quotes per trading day and for the BrokerTec period (2001:01–2007:12) we report the average number of transactions per trading day. The number of Treasury bonds and bills totals to 1005 in our sample period. These were transacted or quoted more than 37.7 million times in the on-the-run secondary market.

Table A-1 Average number of quotes/trades per day in the GovPX and BrokerTec databases

Bond maturity	GovPX period	BrokerTec period
3M	374	–
6M	352	–
2Y	2170	1035
3Y	1385	773
5Y	3128	1987
7Y	637	–
10Y	2649	1956
30Y	793	633

A.2. Testing for microstructure noise

To avoid potential bias in the estimates of the realized volatility using high-frequency data, we apply several tests for the presence of noise caused by the market microstructure effects. In a first step, we compute the first order autocorrelation in high-frequency price returns. Table A-2 reports the first order autocorrelation of equally-spaced ten-minute yield changes in the US Treasury zero curve in the period 1992:01–2007:12. The autocorrelation is statistically significant for the maturities of three, five and ten years. However, the magnitude of all autocorrelations is very small, which makes them economically insignificant.

Table A-2 Autocorrelation of high-frequency yield changes

	2Y	3Y	5Y	7Y	10Y
Autocorrelation	0.0039	0.0154	-0.0064	-0.0040	-0.0175
p-value	(0.0648)	(0.0001)	(0.0025)	(0.0589)	(0.0001)

In a second step, we use the volatility signature plots displaying the average realized volatility against the sampling frequency (Fig. A-1). In the presence of microstructure noise, the average realized volatility increases with the sampling frequency. The reason is the dominance of noise at the very high-frequency sampling (see e.g., Bandi and Russell, 2008). None of the above diagnostics suggests that the microstructure noise present in our data is large and could overwhelm our results.

A.3. Extracting zero coupon yield curve from high-frequency data

We fit the discount curve using smoothing splines. One of the important steps in the procedure is to select the appropriate number of knot points. We make the number of knot points dependent on the number of available bonds and locate them at the bond maturities. In our setting, the number of knot points varies between three and six. The fact that we consider only one specific part of the zero coupon yield curve allows us to use constant roughness penalty as in Fisher, Nychka, and Zervos (1994) for estimating the whole curve. Waggoner (1997) proposes a varying roughness penalty for the smoothing splines procedure with a low penalty at the short end and a high penalty at the

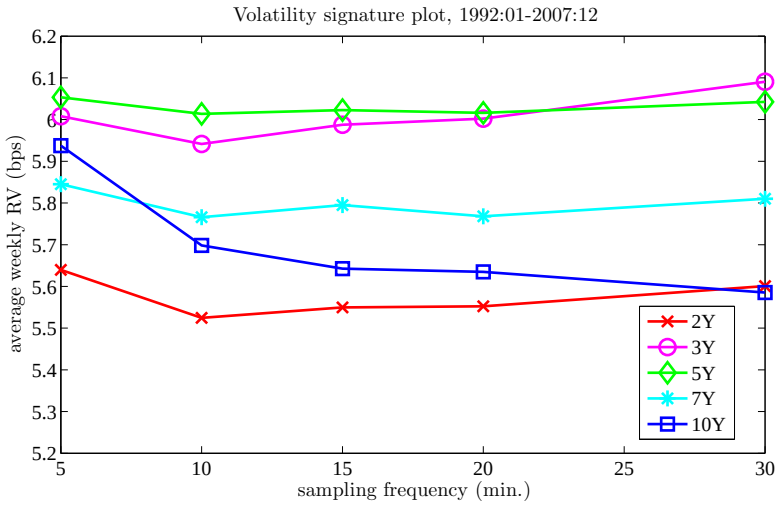


Fig. A-1. Volatility signature plot

We plot the average weekly realized volatility (RV) against the sampling frequency for the whole sample 1992:01–2007:12. We consider five maturities in the zero coupon curve: two, three, five, seven, and ten years.

Table A-3 Correlation of zero coupon yields with CMT and GSW yields

	2Y	3Y	5Y	7Y	10Y
Corr CMT	1.000	1.000	0.999	0.996	0.997
Corr GSW	1.000	0.999	0.999	0.998	0.997

very long end of the curve. In the period 2001:01–2007:12, the intraday quotes on Treasury bills are not available from the BrokerTec database. In order to anchor the very short end for the smoothing splines procedure, we include the daily data on the three-month Treasury bill obtained from the FRED database at the FRB St. Louis. Before using the constructed zero curve for the realized volatility estimation, we compare our zero coupon yields with the daily Constant Maturity Treasury rates (CMT) from the Fed, as well as with zero yields compiled by Gurkaynak, Sack, and Wright (2006) (GSW). Our daily yields are virtually perfectly correlated with the CMTs as well as with the GSW yields. Table A-3 summarizes the results.

A.4. Survey data

BlueChip Financial Forecasts. BlueChip Financial Forecasts (BCFF) survey contains monthly forecast of yields, inflation and GDP growth given by approximately 45 leading financial institutions. The BCFF is published on the first day of each month, but the survey itself is conducted over a two-day period, usually between the 23rd and 27th of each month. The exception is the survey for the January issue which generally takes place between the 17th and 20th of December. The precise dates as to when the survey was conducted are not published. The BCFF provides forecasts of constant maturity yields across several maturities: three and six months, one, two, five, ten, and 30 years. The short end of the term structure is additionally covered with the forecasts of the Fed funds rate, prime bank rate and three-month LIBOR rate. The forecasts are quarterly averages of interest rates for the current quarter, the next quarter out to five quarters ahead. The figures are expressed as percent per annum. In addition, panelist provide forecasts for macroeconomic quantities: real GDP, GDP price index and Consumer Price Index (CPI). The numbers are seasonally adjusted quarter-on-quarter changes.

BlueChip Economic Indicators. The BlueChip Economic Indicators (BCEI) survey contains individual and consensus forecasts of about 50 professional economists from leading financial and advisory institutions. The survey is compiled on a monthly basis, and contains predictions of key financial and macroeconomic indicators, e.g. real and nominal GDP, GDP deflator, CPI, three-month T-bill rate, industrial production, unemployment, housing starts. The survey is conducted over two days, generally beginning on the first business day of each month. The newsletter is typically finished on the third day following completion of the survey and published on the tenth of a month. Every month, panelists provide two types of forecasts: (*i*) average figure for the current calendar year and (*ii*) average

figure for the next calendar year. For instance, in January 2001 the survey contains forecasts for 2001 and 2002. In February 2001, the forecast horizon shrinks to 11 months for the current year, and to 23 months for the next year, and so on. The diminishing forecast horizon implies that the cross-sectional uncertainty measures computed from the individual responses display a visible seasonal pattern. To gauge uncertainty, every month we use the mean absolute deviation of individual forecasts. To remove the problem of seasonality, we adjust the series with a X-12 ARIMA filter. Consensus forecast is defined as the median of individual forecasts in a given month.

Appendix B.

This Appendix derives moments of the state variables necessary for the implementation of the unscented Kalman filter.

B.1. Moments of the V_t process

The first conditional moment of the volatility process V_t is given as:

$$E_t\left(V_{t+\Delta t}\right) = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t}, \tag{37}$$

where

$$\Phi_{V,\Delta t} = e^{M\Delta t} \tag{38}$$

$$\bar{\mu}_{V,\Delta t} = \int_0^{\Delta t} e^{M\Delta t} Q'Q e^{M'\Delta t} ds = -\frac{1}{2}\hat{C}_{12}(\Delta t)\hat{C}'_{11}\left(\Delta t\right), \tag{39}$$

with

$$\begin{pmatrix} \hat{C}_{11}(\Delta t) & \hat{C}_{12}(\Delta t) \\ \hat{C}_{21}(\Delta t) & \hat{C}_{22}(\Delta t) \end{pmatrix} = \exp\left[\Delta t \begin{pmatrix} M & -2Q'Q \\ 0 & -M' \end{pmatrix}\right].$$

Assuming stationarity (i.e. negative eigenvalues of M), the unconditional first moment of V_t follows as:

$$\lim_{\Delta t\to\infty} \text{vec}E_t\left(V_{t+\Delta t}\right) = k\text{vec}\left(\bar{\mu}_{V,\infty}\right) = -k\left[\left(I\otimes M\right) + \left(M\otimes I\right)\right]^{-1}\text{vec}(Q'Q). \tag{40}$$

The conditional and unconditional covariance matrix of V_t reads:

$$Cov_t\left[\text{vec}\left(V_{t+\Delta t}\right)\right] = \left(I_{n^2} + K_{n,n}\right)\left[\Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t}\otimes\bar{\mu}_{V,\Delta t} + k\left(\bar{\mu}_{V,\Delta t}\otimes\bar{\mu}_{V,\Delta t}\right) + \bar{\mu}_{V,\Delta t}\otimes\Phi_{V,\Delta t}V_t\Phi'_{V,\Delta t}\right]. \tag{41}$$

$$\lim_{\Delta t\to\infty} Cov_t\left[\text{vec}\left(V_{t+\Delta t}\right)\right] = \left(I_{n^2} + K_{n,n}\right)k\left(\bar{\mu}_{V,\infty}\otimes\bar{\mu}_{V,\infty}\right). \tag{42}$$

$K_{n,n}$ is the commutation matrix with the property that $K_{n,n}\text{vec}(A) = \text{vec}(A')$. These moments are derived in Buraschi et al. (2008) and thus are stated without a proof.

Gourieroux et al. (2005) show that when $\Omega\Omega' = kQ'Q$, k integer, the dynamics of V_t can be represented as the sum of outer products of k independent Ornstein-Uhlenbeck processes with a zero long-run mean:

$$V_t = \sum_{i=1}^{k} v_t^i v_t^{i\prime} \tag{43}$$

$$v_{t+\Delta t}^i = \Phi_{V,\Delta t} v_t^i + \epsilon_{t+\Delta t}^i, \quad \epsilon_t^i \sim N(0, \bar{\mu}_{V,\Delta t}). \tag{44}$$

Taking the outer-product implies that the exact discretization of V_t has the form:

$$V_{t+\Delta t} = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} + u^V_{t+\Delta t}, \tag{45}$$

where the shock $u^V_{t+\Delta t}$ is a heteroskedastic martingale difference sequence.

B.2. Moments of the Y_t dynamics

We assume that the dimension of X_t is $n = 2$ and f_t is a scalar process. Let $Y_t = (X'_t, f_t)'$:

$$dY_t = \left(\mu_Y + \mathcal{K}_Y Y_t\right)dt + \Sigma\left(V_t\right)dZ_t. \tag{46}$$

It is straightforward to show that the conditional and unconditional first moment of Y_t has the form:

$$E_t\left(Y_{t+\Delta t}\right) = \left(e^{\mathcal{K}_Y\Delta t} - I\right)\mathcal{K}_Y^{-1}\mu_Y + e^{\mathcal{K}_Y\Delta t}Y_t \tag{47}$$

$$\lim_{\Delta t\to\infty} E_t\left(Y_{t+\Delta t}\right) = -\mathcal{K}_Y^{-1}\mu_Y, \tag{48}$$

where $\mathcal{K}_Y$ is assumed to be lower triangular with negative eigenvalues.

To compute the conditional covariance of Y_t, let $V_Y\left(t,T\right) := Cov_t\left(Y_T\right)$. Following Fisher and Gilles (1996), the application of Ito's lemma to $\hat{Y}(t,T) := E_t\left(Y_T\right)$ reveals that:

$$d\hat{Y}\left(t,T\right)=\hat{\sigma}_{Y}\left(t,T\right)dZ_{t}, \tag{49}$$

where $\hat{\sigma}_{Y}(t,T):=\Phi_{Y}\left(t,T\right)\Sigma\left(V_{t}\right)$, with

$$\Phi_{Y}\left(t,T\right)=e^{\mathcal{K}_{Y}(T-t)} \tag{50}$$

and

$$\Sigma_{Y}\left(V_{t}\right)=\left(\begin{array}{cc}\sqrt{V_{t}} & 0\\ 0 & \sigma_{f}^{2}\end{array}\right). \tag{51}$$

Then, integrating $d\hat{Y}\left(t,T\right)$ yields:

$$Y_{T}=\hat{Y}_{T,T}=\hat{Y}_{t,T}+\int_{s=t}^{T}\hat{\sigma}_{Y}\left(s,T\right)dZ_{s}. \tag{52}$$

Therefore, we have:

$$V_{Y}(t,T)=Cov_{t}\left[\int_{s=t}^{T}\hat{\sigma}_{Y}(s,T)dZ_{s}^{Y}\right]=E_{t}\left[\int_{s=t}^{T}\hat{\sigma}_{Y}(s,T)\hat{\sigma}_{Y}(s,T)'ds\right] \tag{53}$$

$$=\int_{s=t}^{T}\Phi_{Y}(s,T)E_{t}\left(\begin{array}{cc}V_{s} & 0\\ 0 & \sigma_{f}^{2}\end{array}\right)\Phi_{Y}'(s,T)ds. \tag{54}$$

Note that since $\mathcal{K}_{Y}$ is lower triangular, $\Phi_{Y}(t,T)=e^{\mathcal{K}_{Y}(T-t)}$ is also lower triangular, and we have:

$$\Phi_{Y}\left(t,T\right)=\left(\begin{array}{cc}\Phi_{X}\left(t,T\right) & 0\\ \Phi_{Xf}\left(t,T\right) & \Phi_{f}\left(t,T\right)\end{array}\right). \tag{55}$$

Let us for convenience define two matrices:

$$\mathcal{M}_{1Y}(t,T)=\left(\begin{array}{c}\Phi_{X}(t,T)\otimes\Phi_{X}(t,T)\\ \Phi_{X}(t,T)\otimes\Phi_{fX}(t,T)\\ \Phi_{fX}(t,T)\otimes\Phi_{X}(t,T)\\ \Phi_{fX}(t,T)\otimes\Phi_{fX}(t,T)\end{array}\right)\text{ and }\mathcal{M}_{0Y}=\left(\begin{array}{c}0_{8\times1}\\ \Phi_{f}^{2}(t,T)\sigma_{f}^{2}\end{array}\right). \tag{56}$$

With help of simple matrix algebra applied to (54), the conditional covariance of Y_t has the (vectorized) form

$$\text{vec}V_{Y}(t,T)=\int_{s=t}^{T}\mathcal{M}_{1Y}(s,T)\left[\Phi_{V}(s,T)\otimes\Phi_{V}(s,T)\right]ds\times\text{vec}\left(V_{t}\right)+\int_{s=t}^{T}k\mathcal{M}_{1Y}\text{vec}\left[\bar{\mu}_{V}(t,s)\right]ds+\int_{s=t}^{T}\mathcal{M}_{0Y}(s,T)ds. \tag{57}$$

The unconditional covariance of Y is given as:

$$\lim_{T\to\infty}\text{vec}V_{Y}(t,T)=\lim_{T\to\infty}\int_{s=t}^{T}k\mathcal{M}_{1Y}(s,T)\text{vec}\left[\bar{\mu}_{V}(t,s)\right]ds+\int_{s=t}^{T}\mathcal{M}_{0Y}(s,T)ds. \tag{58}$$

This expression exists if the mean reversion matrices M and $\mathcal{K}_{Y}$ are negative definite.

The expressions for the conditional mean (47) and covariance (57) give rise to an exact discretization of the process Y_t.

Remark 1. In order to avoid the numerical integration, we can resort to a discrete-time approximation of the unconditional covariance matrix of Y factors. To this end, we discretize the dynamics

$$dY_{t}=\left(\mu_{Y}+\mathcal{K}_{Y}Y_{t}\right)dt+\Sigma_{Y}\left(V_{t}\right)dZ_{t} \tag{59}$$

as

$$Y_{t+\Delta t}=\bar{\mu}_{Y,\Delta t}+\Phi_{Y,\Delta t}Y_{t}+\Sigma_{Y}\left(V_{t}\right)\sqrt{\Delta t}\varepsilon_{t+\Delta t}, \tag{60}$$

where $\bar{\mu}_{Y,\Delta t}=\left(e^{\mathcal{K}_{Y}\Delta t}-I\right)\mathcal{K}_{Y}^{-1}\mu_{Y}$. The second moment of the discretized dynamics is straightforward to obtain as:

$$\begin{aligned}\text{vec}E\left(YY'\right)&=\left(I-\Phi_{Y,\Delta t}\otimes\Phi_{Y,\Delta t}\right)^{-1}\times\\ &\times\text{vec}\left\{\bar{\mu}_{Y,\Delta t}\bar{\mu}_{Y,\Delta t}'+\bar{\mu}_{Y,\Delta t}E\left(Y'\right)\Phi_{Y,\Delta t}'+\Phi_{Y,\Delta t}E\left(Y\right)\bar{\mu}_{Y,\Delta t}'+E\left[\Sigma_{Y}\left(V_{t}\right)\Sigma_{Y}\left(V_{t}\right)'\right]\Delta t\right\}\\ \text{vec}\left[Var\left(Y\right)\right]&=\text{vec}E\left(YY'\right)-\text{vec}E\left(Y\right)\left[\text{vec}E\left(Y\right)\right]'.\end{aligned} \tag{61}$$

We check that for the weekly discretization step $\Delta t=\frac{1}{52}$ this approximation works well, and implies a significant reduction of the computational time.

Appendix C.

We provide a solution for the general version of the model, which incorporates both correlation between the dW and dZ shocks and a general form of the market prices of risk. Based on arguments presented in the body of the paper, we analyze a restricted version of the model, in which the correlation parameter is set to zero and only dZ shocks are priced.

C.1. Dependence between X and V factors

In the general case, X_t and V_t can be correlated, i.e.:

$$dZ_X = dW\rho + \sqrt{1-\rho'\rho}\, dB \tag{62}$$

$$= dW\rho + \widetilde{\rho} dB, \tag{63}$$

where dB is a (2×1)-vector of Brownian motions which is independent from dW, and ρ is a (2×1)-vector such that $\rho \in [-1,1]$ and $\rho'\rho < 1$ (e.g., da Fonseca, Grasselli, and Tebaldi, 2006; Buraschi, Porchia, and Trojani, 2009). We use short notation $\widetilde{\rho} := \sqrt{1-\rho'\rho}$.

C.2. General form of the market prices of risk

Let us write the shocks to Y under the physical dynamics as (for brevity we omit the superscript $\mathbb{P}$):

$$dZ = \begin{pmatrix} dZ_X \\ dZ_f \end{pmatrix} = \begin{pmatrix} dW\rho + \widetilde{\rho} dB \\ dZ_f \end{pmatrix} = \begin{pmatrix} dW\rho \\ 0_{1\times 2} \end{pmatrix} + \underbrace{\begin{pmatrix} \widetilde{\rho} I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & 1_{1\times 1} \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} dB \\ dZ_f \end{pmatrix}}_{d\widetilde{Z}} = \begin{pmatrix} dW\rho \\ 0_{1\times 2} \end{pmatrix} + R d\widetilde{Z}, \tag{64}$$

where

$$R = \begin{pmatrix} \widetilde{\rho} I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & 1_{1\times 1} \end{pmatrix}, \qquad d\widetilde{Z} = \begin{pmatrix} dB \\ dZ_f \end{pmatrix}. \tag{65}$$

The change of drift is specified as:

$$d\widetilde{Z} = d\widetilde{Z}^{\mathbb{Q}} - \Lambda_{Y,t} dt \tag{66}$$

$$dW = dW^{\mathbb{Q}} - \Lambda_{V,t} dt \tag{67}$$

$$\Lambda_{Y,t} = \Sigma_Y^{-1}(V_t)\left(\lambda_Y^0 + \lambda_Y^1 Y_t\right) \tag{68}$$

$$\Lambda_{V,t} = \left(\sqrt{V_t}\right)^{-1} \Lambda_V^0 + \sqrt{V_t}\Lambda_V^1, \tag{69}$$

where λ_Y^0 and λ_Y^1 are a $(n+1)$-vector and $(n+1)\times(n+1)$ matrix of parameters, and Λ_V^0 and Λ_V^1 are $n\times n$ constant matrices. To exclude arbitrage, the market price of risk requires that the parameter matrix Q be invertible, so that V_t stays in the positive-definite domain. This specification implies the risk-neutral dynamics of Y_t given by:

$$dY_t = \left[\left(\mu_Y - \begin{pmatrix} \Lambda_V^0 \rho \\ 0 \end{pmatrix} - R\lambda_Y^0\right) + \left(\mathcal{K}_Y - R\lambda_Y^1\right) Y_t - \begin{pmatrix} V_t \Lambda_V^1 \rho \\ 0 \end{pmatrix}\right] dt + \Sigma_Y\left(V_t\right) dZ_t^{\mathbb{Q}}. \tag{70}$$

Let:

$$\mu_Y^Q = \mu_Y - \begin{pmatrix} \Lambda_V^0 \rho \\ 0 \end{pmatrix} - R\lambda_Y^0 \tag{71}$$

$$\mathcal{K}_Y^Q = \mathcal{K}_Y - R\lambda_Y^1. \tag{72}$$

The dynamics of V_t is given as:

$$dV_t = \left[\left(\Omega\Omega' - \Lambda_V^0 Q - Q'\Lambda_V^{0\prime}\right) + \left(M - Q'\Lambda_V^{1\prime}\right) V_t + V_t\left(M' - \Lambda_V^1 Q\right)\right] dt + \sqrt{V_t} dW_t^{\mathbb{Q}} Q + Q' dW_t^{\mathbb{Q}\prime} \sqrt{V_t}. \tag{73}$$

Let

$$\Omega^{\mathbb{Q}}\Omega^{\mathbb{Q}\prime} = \Omega\Omega' - \Lambda_V^0 Q - Q'\Lambda_V^{0\prime} = (k-2v) Q'Q \tag{74}$$

$$M^{\mathbb{Q}} = M - Q'\Lambda_V^{1\prime}, \tag{75}$$

where, to preserve the same distribution under $\mathbb{P}$ and $\mathbb{Q}$, we assume $\Lambda_V^0 = vQ'$ for a scalar v such that $(k-2v) > n-1$.

C.3. Solution for bond prices

Since both components of the state vector, i.e. Y_t, V_t, are affine, bond prices are of the form:

$$F\left(Y_t, V_t; t, \tau\right) = \exp\left\{A\left(\tau\right) + B\left(\tau\right)' Y_t + Tr\left[C\left(\tau\right) V_t\right]\right\}. \tag{76}$$

By discounted Feynman-Kac theorem, the drift of dF equals rF, thus:

$$\mathcal{L}_{\{Y,V\}} F + \frac{\partial F}{\partial t} = rF, \tag{77}$$

where $\mathcal{L}_{\{Y,V\}}$ is the joint infinitesimal generator of the couple $\{Y_t, V_t\}$ under the risk neutral measure. We have:

$$\mathcal{L}_{\{Y,V\}} F = \left(\mathcal{L}_Y + \mathcal{L}_V + \mathcal{L}_{Y,V}\right) F \tag{78}$$

$$\mathcal{L}_Y F = \frac{\partial F}{\partial Y'}\left[\mu_Y^{\mathbb{Q}} + \mathcal{K}_Y^{\mathbb{Q}} Y - \begin{pmatrix} V_t \Lambda_V^1 \rho \\ 0 \end{pmatrix}\right] + \frac{1}{2} Tr\left[\frac{\partial F}{\partial Y \partial Y'} \Sigma_Y\left(V\right) \Sigma_Y'\left(V\right)\right] \tag{79}$$

$$\mathcal{L}_V F = Tr\left[\left(\Omega^{\mathbb{Q}} \Omega^{\mathbb{Q}\prime} + M^{\mathbb{Q}} V + V M^{\mathbb{Q}\prime}\right) \mathcal{R} F + 2V \mathcal{R} Q' Q \mathcal{R} F\right] \tag{80}$$

$$\mathcal{L}_{Y,V} F = 2Tr\left[\left(\mathcal{R} Q' \rho \frac{\partial}{\partial X'}\right) FV\right]. \tag{81}$$

$\mathcal{R}$ is a matrix differential operator: $\mathcal{R}_{ij} := \left(\frac{\partial}{\partial V_{ij}}\right)$. Substituting derivatives of (76) into (77) gives:

$$B_\tau'\left(\mu_Y^{\mathbb{Q}} + \mathcal{K}_Y^{\mathbb{Q}} Y\right) - Tr\left(\Lambda_V^1 \rho B_{X,\tau}' V\right) + \frac{1}{2} Tr\left(B_{X,\tau} B_{X,\tau}' V\right) + \frac{1}{2} B_{f,\tau}^2 \sigma_f^2 \tag{82}$$

$$+ Tr\left(\Omega^{\mathbb{Q}} \Omega^{\mathbb{Q}\prime} C_\tau\right) + Tr\left[\left(C_\tau M^{\mathbb{Q}} + M^{\mathbb{Q}\prime} C_\tau + 2C_\tau Q' Q C_\tau\right) V\right] \tag{83}$$

$$+ Tr\left[\left(C_\tau Q' \rho B_{X,\tau}' + B_{X,\tau} \rho' Q C_\tau\right) V\right] \tag{84}$$

$$= \frac{\partial A_\tau}{\partial \tau} + \frac{\partial B_\tau}{\partial \tau}' Y + Tr\left(\frac{\partial C_\tau}{\partial \tau} V\right) + \gamma_0 + \gamma_Y' Y \tag{85}$$

By matching coefficients, we obtain the system of equations:

$$\frac{\partial A}{\partial \tau} = B_\tau' \mu_Y^{\mathbb{Q}} + \frac{1}{2} B_{f,\tau}^2 \sigma_f^2 + Tr\left(\Omega^Q \Omega^{Q\prime} C_\tau\right) - \gamma_0 \tag{86}$$

$$\frac{\partial B}{\partial \tau} = \mathcal{K}^{\mathbb{Q}\prime} B_\tau - \gamma_Y \tag{87}$$

$$\frac{\partial C}{\partial \tau} = \frac{1}{2} B_{X,\tau} B_{X,\tau}' + C_\tau\left(M^{\mathbb{Q}} + Q' \rho B_{X,\tau}'\right) + \left(M^{\mathbb{Q}\prime} + B_{X,\tau} \rho' Q\right) C_\tau + 2C_\tau Q' Q C_\tau - \Lambda_V^1 \rho B_{X,\tau}' \tag{88}$$

To obtain the solution provided in the text, set $\rho = 0_{2\times 1}$, $\Lambda_V^0 = 0_{2\times 2}$ and $\Lambda_V^1 = 0_{2\times 2}$.

C.4. Instantaneous volatility of yields

The instantaneous volatility of yields is given as:

$$\frac{1}{dt}\left\langle dy_t^{\tau_1}, dy_t^{\tau_2}\right\rangle = \frac{1}{\tau_1 \tau_2} Tr\left[B_{f,\tau_2} B_{f,\tau_1} \sigma_f^2 + \left(B_{X,\tau_1} B_{X,\tau_2}' + 2C_{\tau_2} Q' \rho B_{X,\tau_1}' + 2C_{\tau_1} Q' \rho B_{X,\tau_2}' + 4C_{\tau_1} Q' Q C_{\tau_2}\right) V_t\right]. \tag{89}$$

Proof. The only term which requires clarification is $B_{\tau_1}' dY_t \times Tr\left[C_{\tau_2} dV_t\right] = B_{X,\tau_1}' dX_t \times Tr\left[C_{\tau_2} dV_t\right]$

$$\begin{aligned} B_{X,\tau_1}' dX_t \times Tr\left[C_{\tau_2} dV_t\right] &= B_{X,\tau_1}' \sqrt{V} dZ_X \times Tr\left[C_{\tau_2}\left(\sqrt{V} dW Q + Q' dW' \sqrt{V}\right)\right] \\ &= B_{X,\tau_1}' \sqrt{V}\left(dW \rho + \widetilde{\rho} dB\right) \times 2Tr\left(Q C_{\tau_2} \sqrt{V} dW\right) \\ &= 2Tr\left(C_{\tau_2} Q' \rho B_{X,\tau_1}' V\right) \end{aligned}$$

Where we use the following fact:

$$Tr\left[C\left(\sqrt{V}dWQ+Q'dW'\sqrt{V}\right)\right]=2Tr\left(QC\sqrt{V}dW\right). \tag{90}$$

□

C.5. Conditional covariance of X and V

We consider the conditional covariance matrix of X and V

$$d\left\langle \begin{pmatrix} X_{t,1} \\ X_{t,2} \\ f_t \end{pmatrix}, \begin{pmatrix} V_{t,11} \\ V_{t,12} \\ V_{t,22} \end{pmatrix} \right\rangle = \begin{pmatrix} d(X_1,V_{11}) & d(X_1,V_{12}) & d(X_1,V_{22}) \\ d(X_2,V_{11}) & d(X_2,V_{12}) & d(X_2,V_{22}) \\ d(f,V_{11}) & d(f,V_{12}) & d(f,V_{22}) \end{pmatrix} \tag{91}$$

The elements of the covariance matrix are given by:

$$d\langle X_k,V_{ij}\rangle=\rho'\left(Q_{:,j}V_{ik}+Q_{:,i}V_{jk}\right), \tag{92}$$

where $Q_{:,j}$ denotes the j-th column of matrix Q.

Proof. The expression follows by simple algebra:

$$\begin{aligned}
\frac{1}{dt}d\langle V_{ij},X_k\rangle &= \left[e_i'\left(\sqrt{V}dWQ\right)e_j+e_i'\left(Q'dW'\sqrt{V}\right)e_j\right]\left(e_k'\sqrt{V}dW\rho\right) \\
&= Tr\left(e_je_i'\sqrt{V}dWQ\right)\times Tr\left(\rho e_k'\sqrt{V}dW\right)+Tr\left(e_je_i'Q'dW'\sqrt{V}\right)\times Tr\left(\rho e_k'\sqrt{V}dW\right) \\
&= \mathrm{vec}\left(\sqrt{V}e_ie_j'Q'\right)'\mathrm{vec}\left(\sqrt{V}e_k\rho'\right)+\mathrm{vec}\left(\sqrt{V}e_je_i'Q'\right)'\mathrm{vec}\left(\sqrt{V}e_k\rho'\right) \\
&= Tr\left(Qe_je_i'Ve_k\rho'\right)+Tr\left(Qe_ie_j'Ve_k\rho'\right) \\
&= \rho'\left(Q_{:,j}V_{ik}+Q_{:,i}V_{jk}\right),
\end{aligned} \tag{93}$$

where e_i is the i-th column of the identity matrix. □

C.6. Discrete approximation to the unconditional covariance matrix of X and V

We can use the discretized dynamics of X and V to compute the unconditional covariance matrix:

$$X_{t+\Delta t}=\bar{\mu}_{X,\Delta t}+\Phi_{X,\Delta t}X_t+\sqrt{V_t\Delta t}\left(U_{t+\Delta t}\rho+\widetilde{\rho}b_{t+\Delta t}\right) \tag{94}$$

$$V_{t+\Delta t}=k\bar{\mu}_{V,\Delta t}+\Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}'+\sqrt{V_t\Delta t}U_{t+\Delta t}Q+Q'U_{t+\Delta t}'\sqrt{V_t\Delta t}, \tag{95}$$

where U_t is a 2×2 matrix of Gaussian shocks, and b_t is a 2-vector of Gaussian shocks. The covariance between X and V is computed as :

$$Cov\left[X,\mathrm{vec}(V)\right]=E\left[X\mathrm{vec}\left(V\right)'\right]-E\left(X\right)E\left[\mathrm{vec}\left(V\right)'\right]. \tag{96}$$

The element $E\left[X\mathrm{vec}\left(V\right)'\right]$ reads:

$$\mathrm{vec}E\left[X\left(\mathrm{vec}V\right)'\right]=\left[I_{n^3}-\left(\Phi_V\otimes\Phi_V\right)\otimes\Phi_X\right]^{-1}\left(\mathrm{vec}A+\mathrm{vec}B\right), \tag{97}$$

where A is given as:

$$A=\bar{\mu}_X\mathrm{vec}\left(k\bar{\mu}_V\right)'+\bar{\mu}_X\mathrm{vec}\left[\Phi_VE\left(V_t\right)\Phi_V'\right]+\Phi_XE(X_t)\mathrm{vec}(k\bar{\mu}_V)', \tag{98}$$

and the elemtent (k,ij) of matrix B, associated with the covariance of X_k and V_{ij} has the form:

$$B_{k,ij}=\rho'\left(Q_{:,j}V_{ik}+Q_{:,i}V_{jk}\right)\Delta t, \tag{99}$$

where $B=\begin{pmatrix} B_{1,11} & B_{1,12} & B_{1,21} & B_{1,22} \\ B_{2,11} & B_{2,12} & B_{2,21} & B_{2,22} \end{pmatrix}$. Note that the second and third columns of B are identical.

Appendix D.

This Appendix summarizes the algorithm for the unscented Kalman filtering.

We recast the transition and measurement equations above into one state space. The compound transition equation is given by:

$$S_{t+\Delta t} = A + BS_t + \varepsilon_{t+\Delta t}, \tag{100}$$

and the compound measurement equation is given by:

$$m_t = h(S_t; \Theta) + \vartheta_t. \tag{101}$$

$S_t = (Y_t', \bar{V}_t')'$ and A are $(n + \bar{n} + 1) \times 1$-dimensional vectors, A is given by:

$$A = \begin{pmatrix} (\Phi_{Y,\Delta t} - I)\,\mathcal{K}_Y^{-1}\mu_Y \\ k \cdot \mathcal{E}_n vec(\bar{\mu}_{V,\Delta t}) \end{pmatrix}. \tag{102}$$

B is a block-diagonal matrix of the form:

$$B = \begin{pmatrix} \Phi_{Y,\Delta t} & \mathbf{0}_{n\times\bar{n}} \\ \mathbf{0}_{\bar{n}\times n} & \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n \end{pmatrix}. \tag{103}$$

The vector shocks is of the form:

$$\varepsilon_{t+\Delta t} = \begin{pmatrix} u^Y_{t+\Delta t} \\ \mathcal{E}_n \mathrm{vec}(u^V_{t+\Delta t}) \end{pmatrix}, \tag{104}$$

and its covariance matrix is given by a block-diagonal matrix:

$$Cov_t(\varepsilon_{t+\Delta t}) = \begin{pmatrix} Cov_t(Y_{t+\Delta t}) & \mathbf{0}_{n\times\bar{n}} \\ \mathbf{0}_{\bar{n}\times n} & Cov_t(\bar{V}_{t+\Delta t}) \end{pmatrix}. \tag{105}$$

m_t is a vector of observed yields and volatility measures given by $m_t = (y_t^\tau, v_t^{\tau_i,\tau_j})'$. Model implied yields and volatilities are affine in the state vector. Function $h(\cdot)$ translates the state variables to model implied yields and volatilities:

$$h(S_t; \Theta) = \begin{pmatrix} f(S_t; \Theta) \\ g(V_t; \Theta) \end{pmatrix}. \tag{106}$$

The vector of measurement errors:

$$\vartheta_t = \begin{pmatrix} \sqrt{R_y} e_t^y \\ \sqrt{R_v} e_t^v \end{pmatrix} \tag{107}$$

is Gaussian with the covariance matrix, for six yields and three volatility measurements, is given by:

$$Cov(\vartheta_t) = \begin{pmatrix} \sigma_y^2 \mathbf{I}_6 & \mathbf{0}_{6\times 3} \\ \mathbf{0}_{3\times 6} & diag(\sigma_{i,v}^2)_{i=1,2,3} \end{pmatrix}. \tag{108}$$

The core of UKF is the unscented transformation which approximates a distribution of a nonlinear transformation of any random variable by a set of sample points. In the UKF framework, we apply the unscented transformation recursively to B and $h(\cdot)$.

We define $L_S := n + \bar{n} + 1$. Assume that we know the mean $\bar{S}$ and the covariance P_S of S_t at each point in time t. We form a matrix $\mathcal{S}$ of $2L_S + 1$ sigma vectors:

$$\mathcal{S}_0 = \bar{S} \tag{109}$$

$$\mathcal{S}_i = \bar{S} + \left(\sqrt{(L_S + \lambda)P_S}\right)_i, i = 1, ..., L_S \tag{110}$$

$$\mathcal{S}_i = \bar{S} - \left(\sqrt{(L_S + \lambda)P_S}\right)_{i-L_S}, i = L_S + 1, ..., 2L_S, \tag{111}$$

where $\lambda = \alpha^2(L_S - \kappa) - L_S$ is a scaling parameter governing the spread of sigma points around the mean and $\left(\sqrt{(L_S + \lambda)P_S}\right)_i$ is the i-th column of matrix P_S. Sigma points $\mathcal{S}$ are propagated through function $h(\cdot)$ to get $\mathcal{M}$. The first two moments of m_t are approximated by:

$$\bar{m} \approx \sum_{i=0}^{2L_S} W_i^{\mu} \mathcal{M}_i \tag{112}$$

$$P_S \approx \sum_{i=0}^{2L_S} W_i^{\sigma} (\mathcal{M}_i - \bar{m})(\mathcal{M}_i - \bar{m})', \tag{113}$$

where W^{μ} and W^{σ} denote weights for the mean and the covariance matrix, respectively and are defined as:

$$W_0^{\mu} = \frac{\lambda}{L_S + \lambda} \tag{114}$$

$$W_0^{\sigma} = \frac{\lambda}{L_S + \lambda} + 1 - \alpha^2 + \beta, i = 1, ..., L_S \tag{115}$$

$$W_i^{\mu} = W_i^{\sigma} = \frac{\lambda}{2(L_S + \lambda)}, i = L_S + 1, ..., 2L_S. \tag{116}$$

Parameters α and β, mainly determine higher moments of the distribution.

The UKF Algorithm

1. Initialize at unconditional moments:[23]

$$\hat{S}_0 = \mathbb{E}[S_0] \tag{117}$$

$$P_{S0} = \mathbb{E}[(S_0 - \hat{S}_0)(S_0 - \hat{S}_0)'] \tag{118}$$

   for $k \in 1, ..., \infty$:

2. Compute the sigma points:

$$\mathcal{S}_{k-1} = \left[\begin{array}{ccc} \hat{S}_{k-1} & \hat{S}_{k-1} + \sqrt{(L_S + \lambda) P_{S,k-1}} & \hat{S}_{k-1} - \sqrt{(L_S + \lambda) P_{S,k-1}} \end{array} \right] \tag{119}$$

3. Time update:

$$\mathcal{S}_{k|k-1}^a = B(\mathcal{S}_{k-1}) \tag{120}$$

$$\hat{S}_k^- = \sum_{i=0}^{2L_S} W_i^{\mu} \mathcal{S}_{k|k-1}^a \tag{121}$$

$$P_{Sk}^- = \sum_{i=0}^{2L_S} W_i^{\sigma} (\mathcal{S}_{ik|k-1}^a - \hat{S}_k^-)(\mathcal{S}_{ik|k-1}^a - \hat{S}_k^-)' + Cov_t(\varepsilon_{t+\Delta t}) \tag{122}$$

4. Augment sigma points:

$$\mathcal{S}_{k|k-1} = \left[\begin{array}{ccc} \mathcal{S}_{k|k-1}^a & \mathcal{S}_{0k|k-1}^a + \sqrt{(L_S + \lambda) Cov_t(\varepsilon_{t+\Delta t})} & \mathcal{S}_{0k|k-1}^a - \sqrt{(L_S + \lambda) Cov_t(\varepsilon_{t+\Delta t})} \end{array} \right] \tag{123}$$

$$\mathcal{M}_{k|k-1} = h(\mathcal{S}_{k|k-1}) \tag{124}$$

$$\hat{m}_k^- = \sum_{i=1}^{2L_S} W_i^{\sigma} \mathcal{M}_{i,k|k-1} \tag{125}$$

5. Measurement equations update:

[23] We borrow the algorithm from Wan and van der Merwe (2001).

$$P_{m_k}^{-} = \sum_{i=0}^{2L_S} W_i^{\sigma} (\mathcal{M}_{ik|k-1} - \hat{m}_k^{-})(\mathcal{M}_{ik|k-1} - \hat{m}_k^{-})' + Cov_t(\vartheta_{t+\Delta t}) \tag{126}$$

$$P_{S_k m_k} = \sum_{i=0}^{2L_S} W_i^{\sigma} (\mathcal{S}_{ik|k-1} - \hat{S}_k^{-})(\mathcal{M}_{ik|k-1} - \hat{m}_k^{-})' \tag{127}$$

$$\mathcal{K}_k = P_{S_k m_k} P_{m_k}^{-1} \tag{128}$$

$$\hat{F}_k = \hat{F}_k^{-} + \mathcal{K}_k (m_k - \hat{m}_k^{-}) \tag{129}$$

$$P_k = P_k^{-} - \mathcal{K}_k P_{m_k}^{-} \mathcal{K}_k'. \tag{130}$$

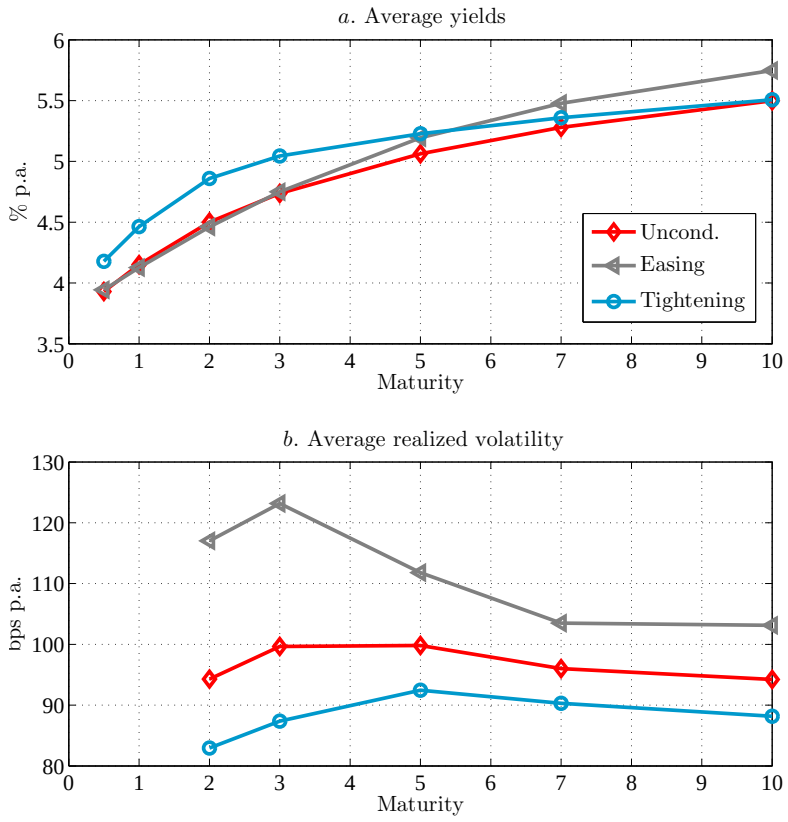


Fig. 1. Yield and volatility curves

The figure plots the yield and volatility curves: unconditional mean for the whole sample 1992:01–2007:12, and means conditional on the Fed's tightening and easing cycles. The cycles are regarded as easing or tightening if at least three subsequent moves in the federal funds rate target have been in the respective direction. In our sample, out of total 3955 days, we identify 646 days as the easing regime, and 958 days as the tightening regime. The realized volatility curves are computed from the daily data and annualized (×250).

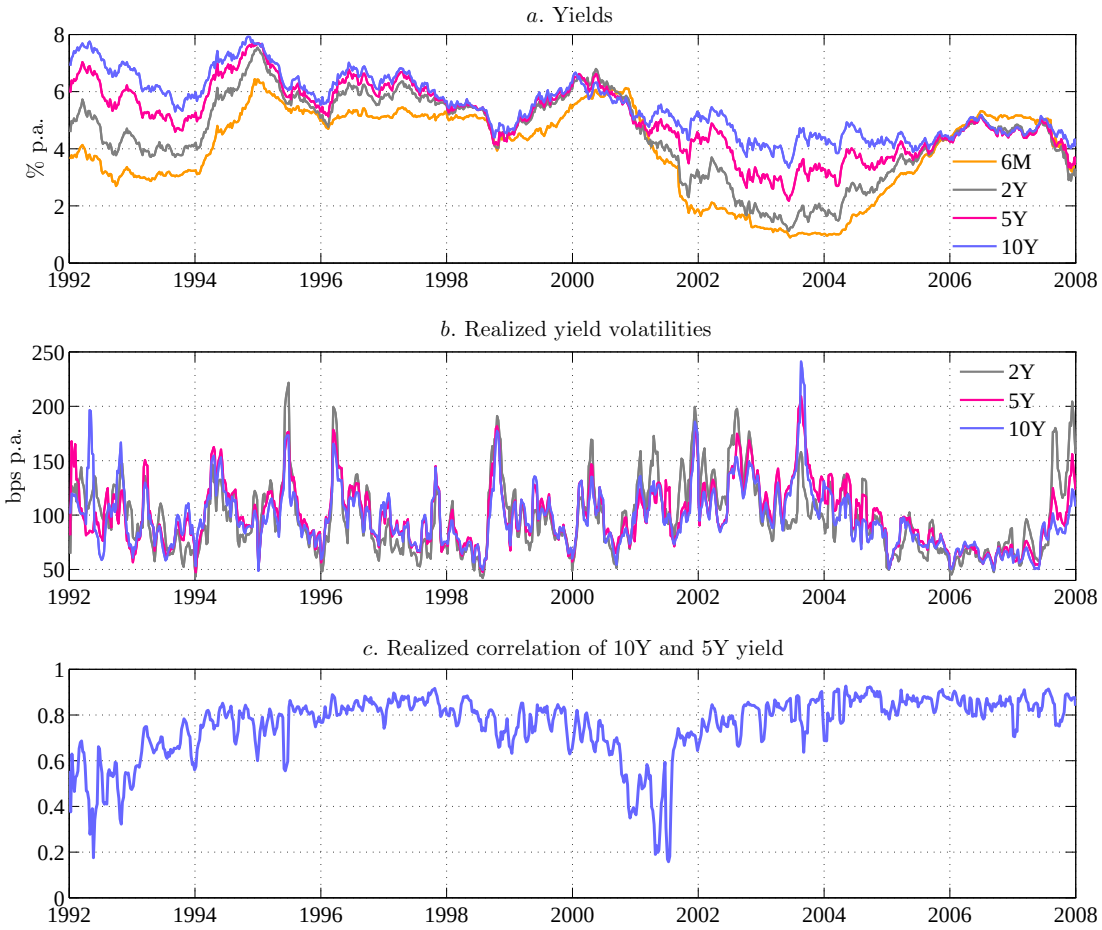


Fig. 2. Evolution of yields and volatilities

The figure plots the dynamics of weekly yields, realized volatilities and correlation over the 1992:01–2007:12 period. Yields include maturities of six months and two, five, and ten years. Realized volatilities are constructed from the actively traded bonds of two, five and ten years to maturity.

Fig. 3. Realized correlations of yield factors

The figure plots realized correlations between yield curve factors (left-hand axis) constructed from the high-frequency zero curve. Level, slope and curvature factor loadings are obtained from the PCA decomposition of the unconditional covariance matrix of yields with maturities: two, three, five and ten years. Daily realized correlations are smoothed over a three-month window. In each panel, correlation dynamics are juxtaposed against the Fed regimes: cut, no change, hike (right-hand axis).

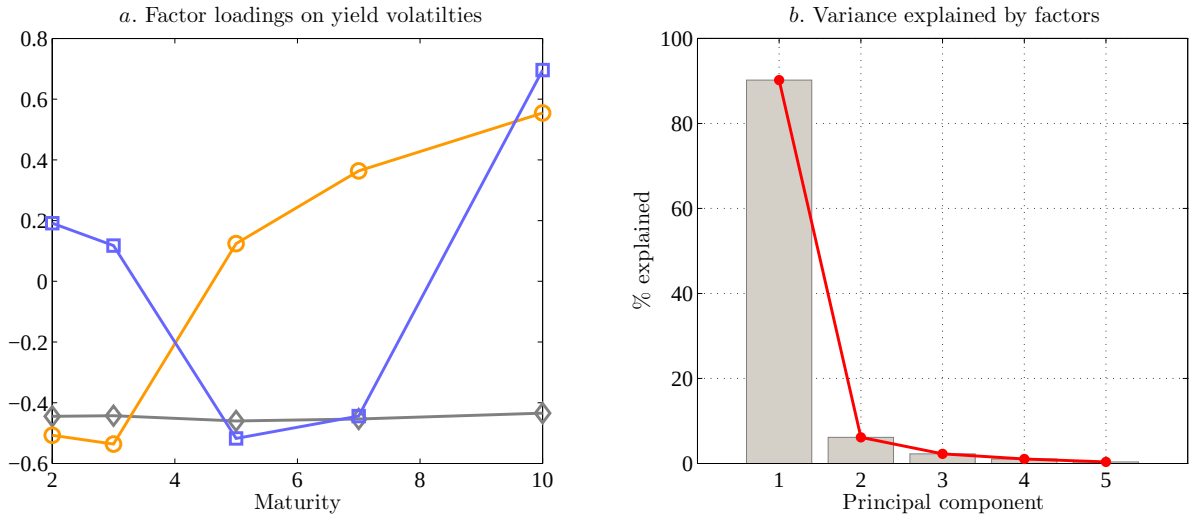


Fig. 4. Factors in yield volatilities

The figure shows the principal component decomposition of the yield volatility curve. We use the unconditional correlation matrix of realized weekly volatilities computed for two-, three-, five-, seven- and ten-year zero bonds. Panel *a* plots the loadings of factors on volatilities. Panel *b* displays the percentage of variance explained by each factor.

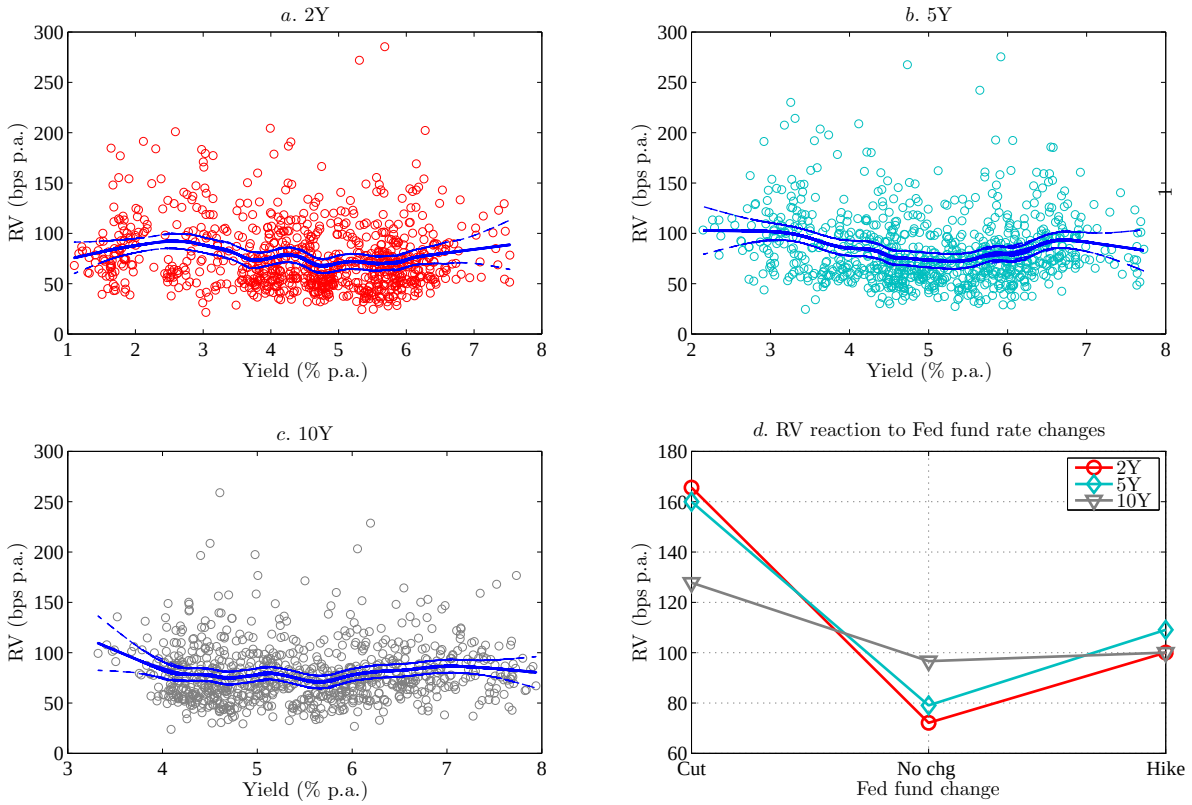


Fig. 5. Yield level versus yield volatility

The figure depicts the relationship between yield level and yield volatility for bonds with two-, five- and ten-year maturity (based on weekly data). In panels *a*–*c*, circles denote data points, and the lines represent the fits of the nonparametric kernel regression together with the 99% confidence bound. Panel *d* shows the level of the realized volatility conditional on the change in the Fed funds target rate (cut, no change, hike). During our sample period 1992:02–2007:12 there were 23 rate cuts, 31 hikes, and 3901 days on which the rate did not change.

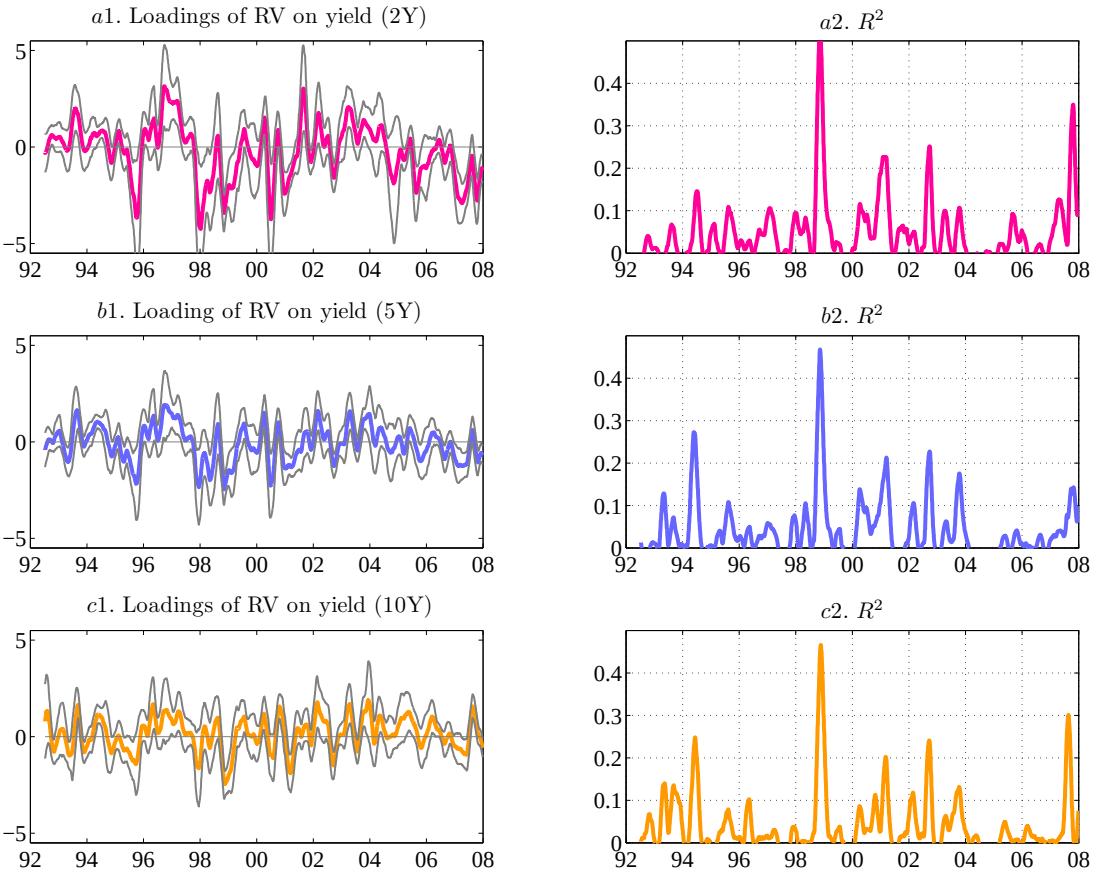


Fig. 6. Regressions of realized volatility on the yield level

The figure plots the rolling regression coefficients, 90% confidence bands (left panels) and R^2's (right panels) in regressions of realized volatilities on yields of corresponding maturities. The regressions use daily data on yields and realized volatilities rolled over the window of 3 months. For clarity, all regression statistics have additionally been smoothed using a window of 50 observations. All variables have been standardized.

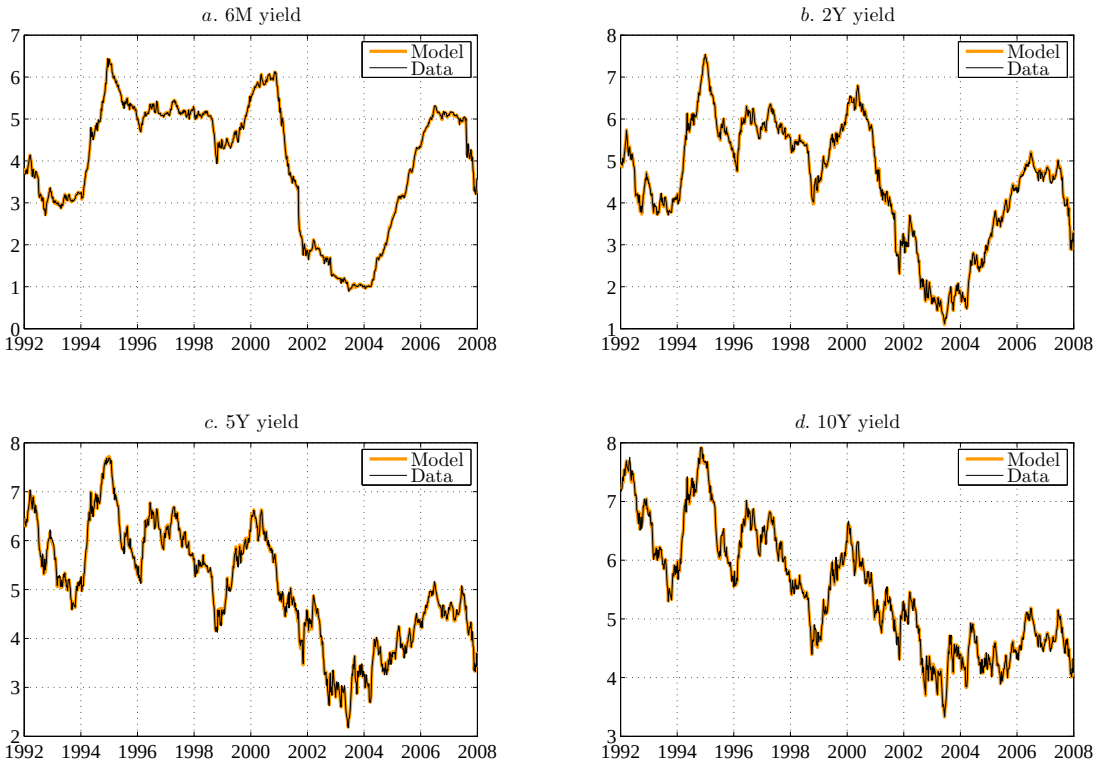


Fig. 7. In-sample fit to yields

The figure compares the in-sample fit of the model-implied yields to the observed yields with maturities: six months, two, five, and ten years over the period 1992:01–2007:12. Yields are presented in percentage p.a., and are measured on weekly frequency.

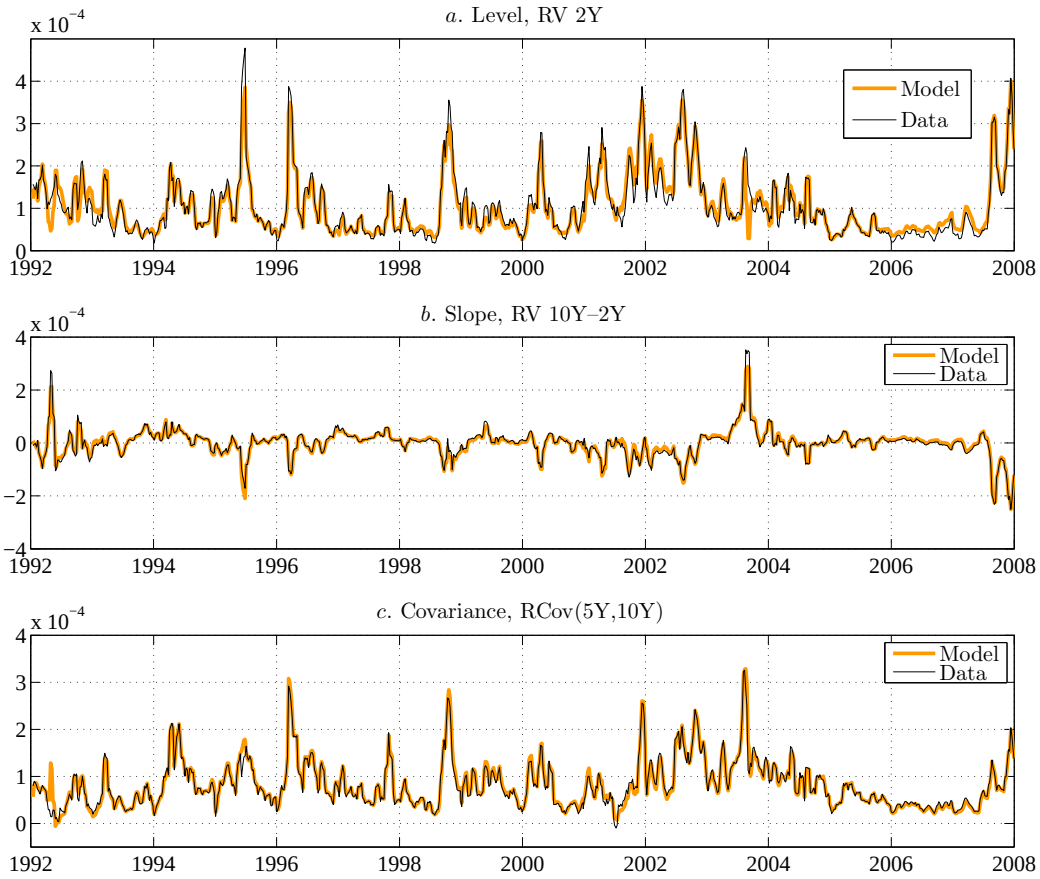


Fig. 8. In-sample fit to the second moment dynamics of yields

The figure plots the model-implied in-sample fit to the yield second moments. The level is defined as the variance of the two-year yield. The slope is computed as the difference between the realized variance of the ten- and the two-year yield. Finally, the covariance is between the five- and the ten-year yield.

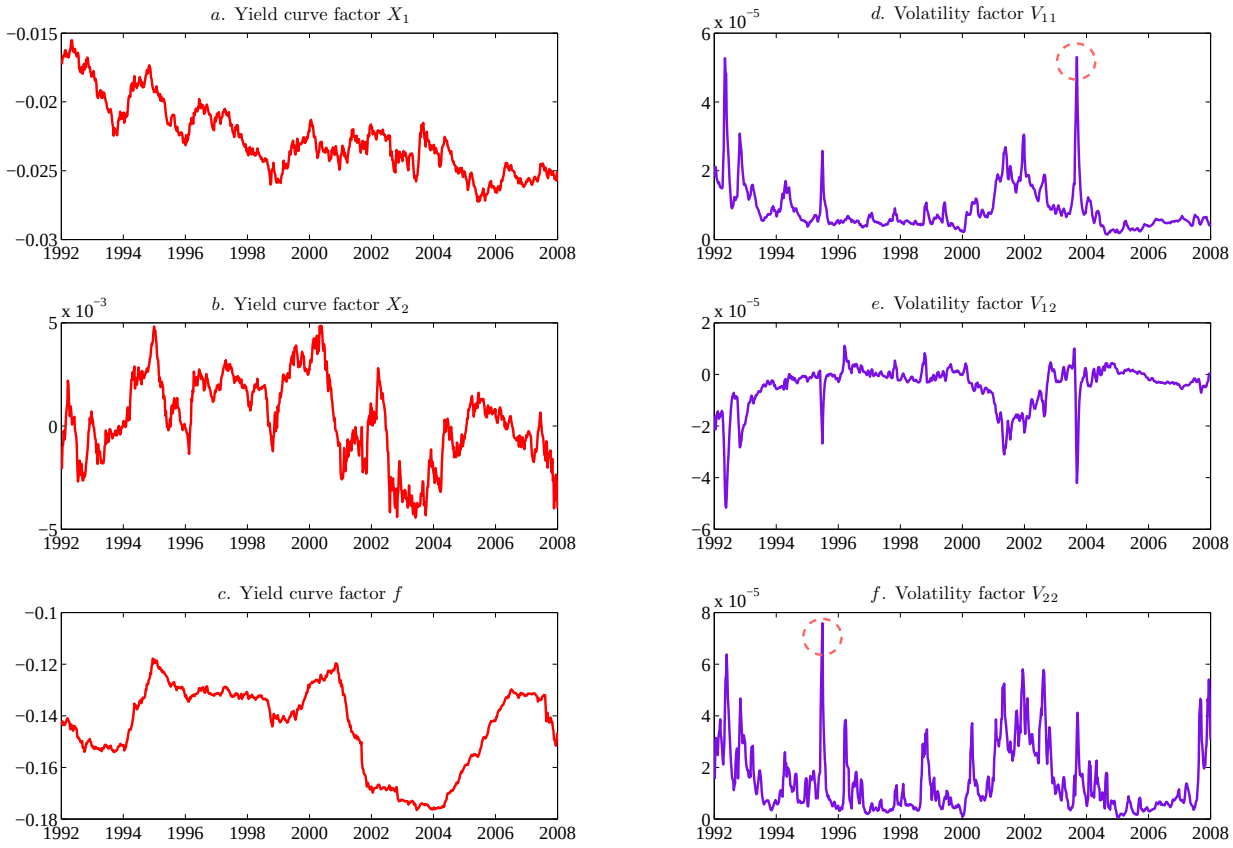


Fig. 9. Filtered factor dynamics

The figure plots the model-implied factor dynamics extracted by the unscented Kalman filter under the estimated parameters of the G_3SV_3 model. The left-hand panels present factors driving the yield curve. The right-hand panels display factors generating time-variation in conditional second moments. The circles in panels d and f mark examples of important events that had a strong impact on the evolution of the volatility curve in our sample period. We discuss them in Section 5.2. For presentation, the X states have been multiplied by -1.

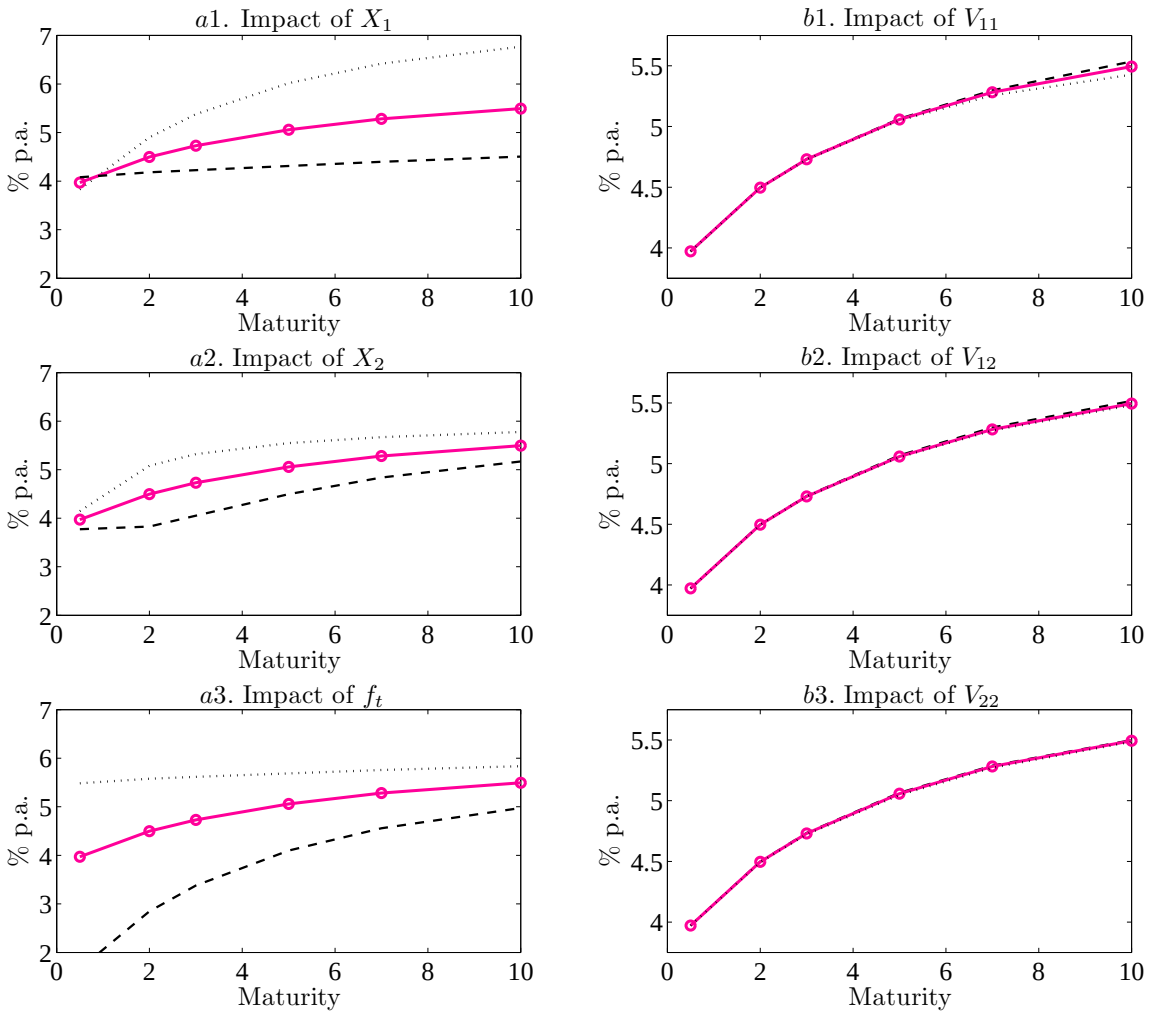


Fig. 10. Factor shocks and yield curve responses

The graph presents the response of the yield curve to shocks in the state variables. The left-hand panels $(a1, a2, a3)$ display the effect of the yield curve factors $Y_t = (X_{1t}, X_{2t}, f_t)'$, the panels on the right $(b1, b2, b3)$ display the effects of the volatility factors $\bar{V}_t = (V_{11t}, V_{12t}, V_{22t})'$. In each subplot, the solid line shows the yield curve generated by setting all state variables to their unconditional means. Circles indicate the maturities used in estimation, i.e. six months, two, three, five, seven, and ten years. The dashed and dotted lines are obtained by setting a given state variable to its tenth and 90th percentile, respectively, and holding the remaining factors at their unconditional average. For presentation, factors X_1, X_2 have been multiplied by -1 so that both correlate positively with yields.

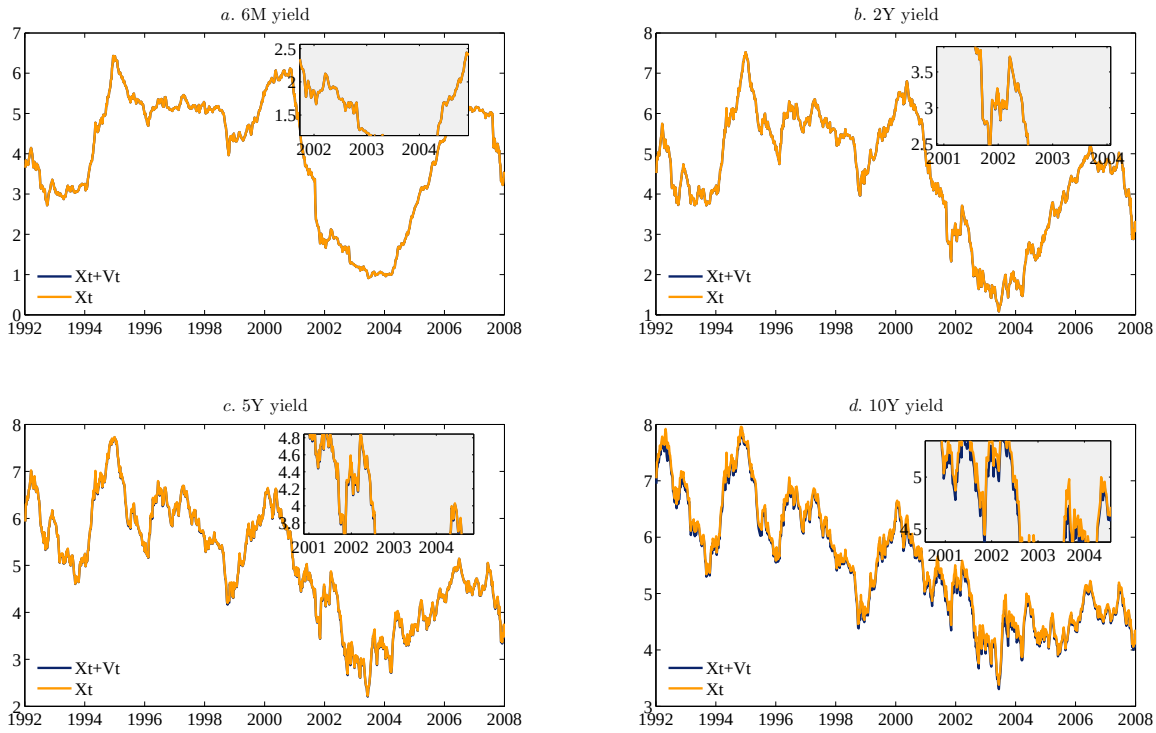


Fig. 11. Contribution of X_t and V_t to model-implied yields

The figure plots the model-implied yields for maturities of six months and two, five and ten years. Each panel compares hypothetical yields that would be generated by X_t factors after V_t factors have been excluded ($y_t^\tau = -\frac{1}{\tau}\{A(\tau) + B(\tau)'X_t\}$) against those produced by X_t and V_t jointly ($y_t^\tau = -\frac{1}{\tau}\{A(\tau)+B(\tau)'X_t+Tr[C(\tau)V_t]\}$). For the ease of comparison, we magnify selected regions of each graph.

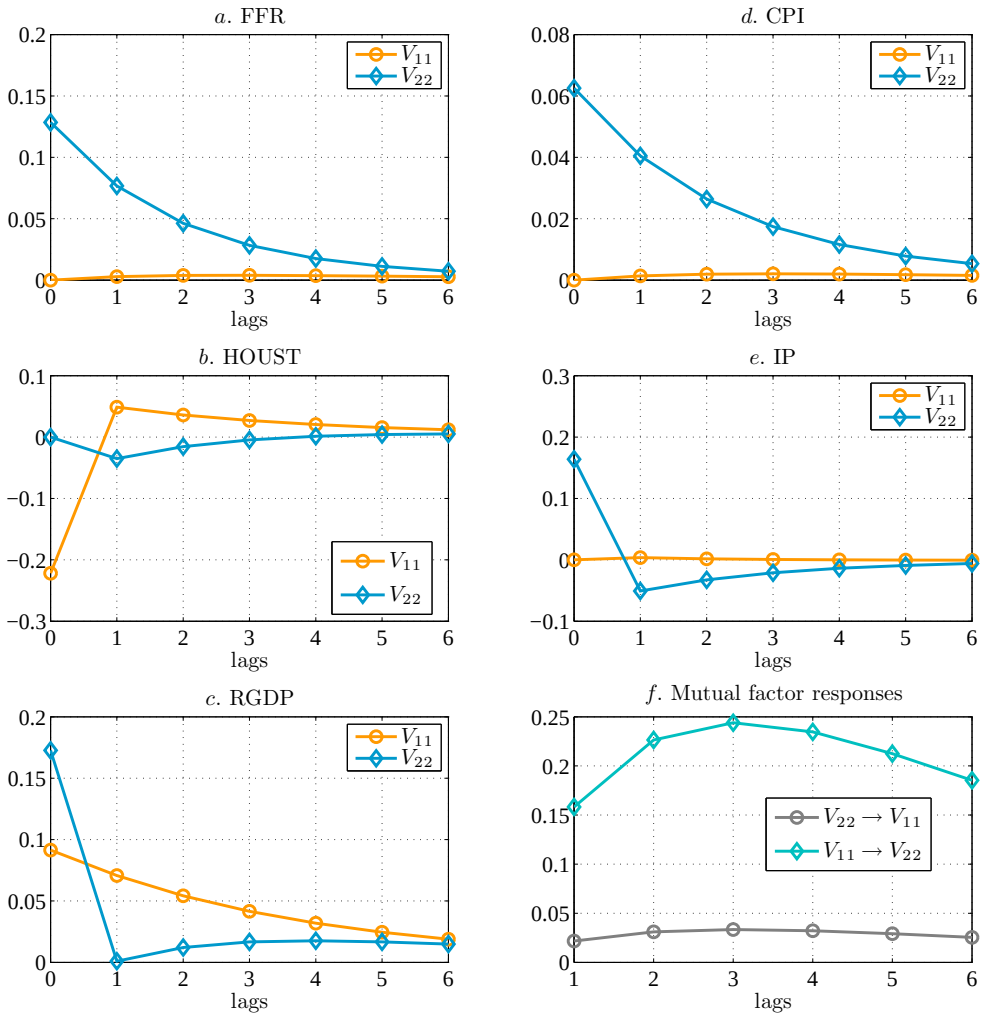


Fig. 12. Impulse-response functions of volatility factors V_{11} and V_{22} to macroeconomic uncertainty proxies

Panels *a*–*e* display the impulse-response functions of volatility factors to macroeconomic uncertainties extracted from the surveys. Uncertainty is proxied with the mean absolute deviation of forecasts made by individual panelists. Panel *f* shows mutual responses of factors to each other. The impulse responses are based on a VARMAX model with first order AR component, allowing for contemporaneous and lagged effects of exogenous macro variables. The impulse responses exclude insignificant exogenous regressors. All variables have been standardized; lags are in months.

Table 1 Major moves in the U.S. Treasury yield and volatility curves in the period 2000–2004 (1st part)

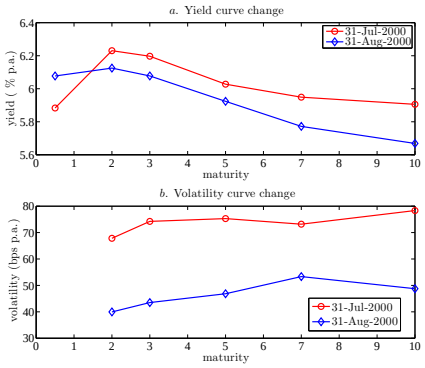
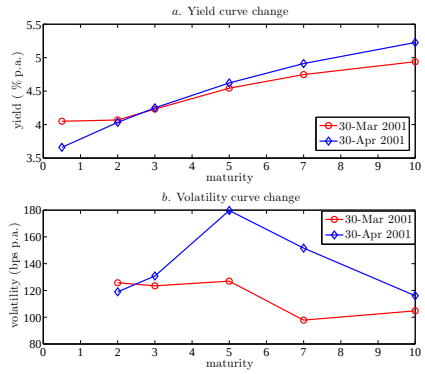
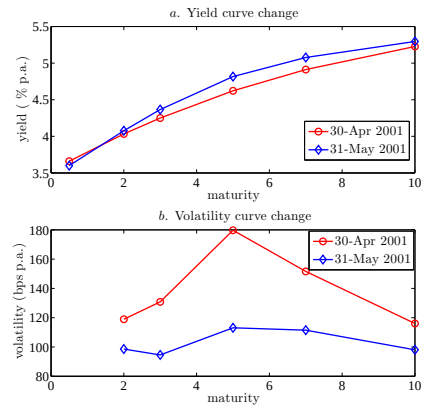
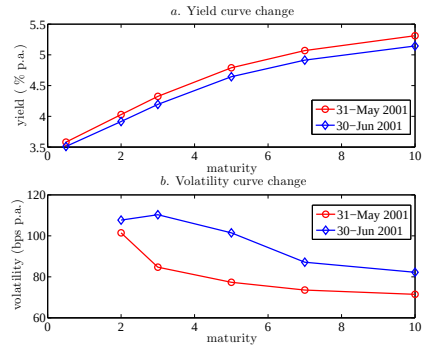
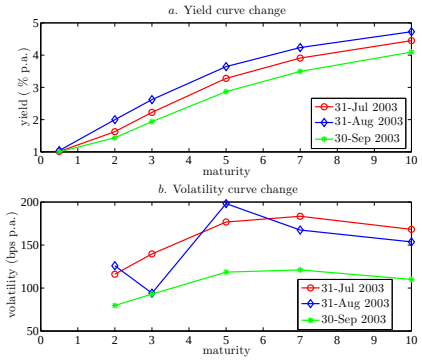
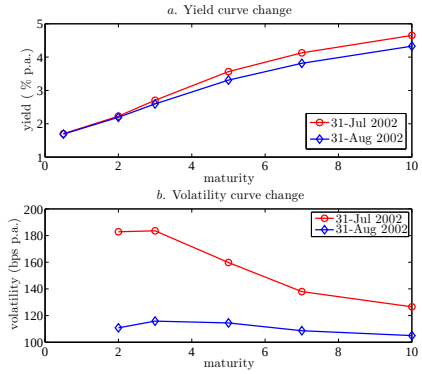
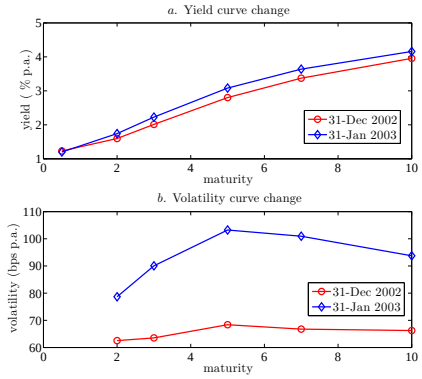
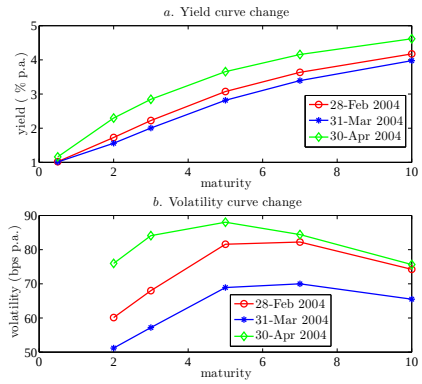
31-Jul-2000, 31-Aug-2000. The main topics are the labor productivity, level of NAIRU and change in the trend of potential output of the US economy. Market participants expect that the increased productivity should keep the inflation lower. At the same time, the economy is slowing from the record GDP growth. All these facts contribute to the rapid decline in the long term interest rates.	
30-Mar-2001, 30-Apr-2001. Leading indicators point to a recession. This leads to the repricing of the short end of the curve as markets expect the Fed to cut interest rates significantly even after the surprising -50 bps move in April 2001.	
30-Apr-2001, 31-May-2001. The flow of disappointing data starts already in April: payrolls, unemployment, all pointing toward recession. Long term yields increase due the concerns that the Fed might be easing too much. In the context of a general market belief: the unemployment data lags the business cycle. The volatility curve moves in the opposite direction to the yield curve and gets hump-shaped.	
31-May-2001, 30-Jun-2001. Recession is being priced-in, as the Fed cuts the rate only by 25 bps instead of 50 bps expected. The volatility increases significantly and changed its shape.	

Table 1 Major moves in the US Treasury yield and volatility curves in the period 2000–2004 (2nd part)

31-Jul-2003, 31-Aug-2003, 30-Sep-2003. The upward shift in yields is caused by the strong labor productivity data, whereas the subsequent downward shift in September is due to the weak August employment data. The FOMC minutes from August meeting are published on September 19th, FOMC members stress the output gap. This leads market participants (together with the worsening data) to revise their expectations.	
31-Jul-2002, 31-Aug-2002. The data flow worsens: ISM manufacturing survey and the unemployment show alarming signals.	
31-Dec-2002, 31-Jan-2003. The ISM data read better than expected, the new orders index increases to 63.3% from 49.9%—the third biggest monthly gain since data is available. The positive news on ISM takes market participants by surprise, this in turn increases the volatility substantially.	
28-Feb-2004, 31-Mar-2004, 30-Apr-2004. The first downward shift in yields is caused by the disappointing data on jobless claims and retail sales. The upward move in yields in April is due to the improved economic data (especially manufacturing), but markets still do not expect Fed's tightening.	

Sources: Bloomberg, Reuters.

Table 2 Descriptive statistics of weekly yields and realized volatilities

The table contains descriptive statistics of weekly yields (panel *a*) and realized yield volatilities (panel *b*) based on the period 1992:01–2007:12. JB denotes the Jarque-Bera normality test (critical value 5.91). Panel *c* shows unconditional correlations between yields and volatilities. Numbers in italics indicate correlations which are not significant at the 1% level.

Panel *a*. Yields (% p.a.)

	6M	2Y	3Y	5Y	7Y	10Y
Mean	3.97	4.50	4.73	5.06	5.28	5.50
Skew	-0.60	-0.46	-0.35	-0.12	0.09	0.31
Kurt	2.15	2.46	2.47	2.30	2.17	2.13
JB	76.22	40.62	27.69	19.08	25.21	40.03

Panel *b*. Realized yield volatilities (bps p.a.)

	2Y	3Y	5Y	7Y	10Y
Mean	108.33	114.44	113.31	108.35	104.62
Stdev	46.73	49.75	45.87	40.75	38.52
Skew	1.47	1.57	1.40	1.18	1.72
Kurt	6.41	7.23	5.98	4.78	9.17
JB	712.90	977.33	589.98	309.12	1755.34

Panel *c*. Unconditional correlations of yields and realized volatilities

	y_t^{6M}	y_t^{2Y}	y_t^{3Y}	y_t^{5Y}	y_t^{7Y}	y_t^{10Y}	v_t^{2Y}	v_t^{3Y}	v_t^{5Y}	v_t^{7Y}	v_t^{10Y}
y_t^{6M}	1										
y_t^{2Y}	0.94	1									
y_t^{3Y}	0.89	0.99	1								
y_t^{5Y}	0.77	0.93	0.97	1							
y_t^{7Y}	0.66	0.85	0.91	0.99	1						
y_t^{10Y}	0.54	0.76	0.84	0.95	0.99	1					
v_t^{2Y}	-0.21	-0.18	-0.15	*-0.10*	*-0.05*	*-0.01*	1				
v_t^{3Y}	-0.24	-0.18	-0.15	*-0.07*	*-0.01*	*0.05*	0.96	1			
v_t^{5Y}	-0.27	-0.22	-0.20	-0.14	*-0.09*	*-0.04*	0.90	0.89	1		
v_t^{7Y}	-0.29	-0.23	-0.20	-0.13	*-0.08*	*-0.02*	0.85	0.86	0.97	1	
v_t^{10Y}	-0.26	-0.19	-0.15	*-0.08*	*-0.02*	*0.03*	0.77	0.78	0.86	0.89	1

Table 3 Parameter estimates

This table reports parameter estimates for four model specifications. "G_3SV_0 risk-neutral" is a Gaussian three-factor model without risk premiums. "G_3SV_0" denotes a Gaussian three-factor model with essentially affine market prices of risk. "G_3SV_3 risk-neutral" is our six-factor model without risk premiums with three conditionally Gaussian factors for the yield curve and three volatility factors. Model specification "G_3SV_3" additionally contains affine risk compensations. The purely Gaussian specifications follow Duffee (2002), and are diffusion-normalized according to the identification scheme of Dai and Singleton (2000). This type of normalization allows to treat γ_Y as a free parameter vector. BHHH standard errors are in parentheses. The last section of the table shows the log-likelihood values. For comparison, in models with stochastic volatility, we split the log-likelihood values into the yield component and the volatility component: Loglik y_t^τ and Loglik v_t^τ.

	G_3SV_0 risk–neutral	G_3SV_0	G_3SV_3 risk–neutral	G_3SV_3
$\mathcal{K}_{Y,11}$	-0.001	-0.506	-0.014	-0.033
	(0.001)	(0.289)	(0.001)	(0.008)
$\mathcal{K}_{Y,21}$	-0.181	-1.653	–	–
	(0.151)	(0.077)	–	–
$\mathcal{K}_{Y,22}$	-0.786	-0.434	-1.233	-1.947
	(0.266)	(0.103)	(0.032)	(0.309)
$\mathcal{K}_{Y,31}$	-1.696	-2.900	-2.797	-2.768
	(0.115)	(0.116)	(0.182)	(0.175)
$\mathcal{K}_{Y,32}$	-2.828	-0.800	-7.275	-7.668
	(0.096)	(0.034)	(0.433)	(0.479)
$\mathcal{K}_{Y,33}$	-0.786	-0.021	-0.500	-0.498
	(0.267)	(0.001)	(0.009)	(0.008)
σ_f	–	–	0.006	0.006
	–	–	(6.96e-5)	(7.05e-5)
k	–	–	2.000	2.000
	–	–	(1.335)	(1.067)
Q_{11}	–	–	0.001	0.001
	–	–	(1.14e-4)	(1.14e-4)
Q_{22}	–	–	0.004	0.003
	–	–	(2.68e-4)	(2.63e-4)
M_{11}	–	–	-0.069	-0.113
	–	–	(0.044)	(0.038)
M_{21}	–	–	-0.260	-0.185
	–	–	(0.184)	(0.177)
M_{22}	–	–	-0.518	-0.685
	–	–	(0.090)	(0.127)
γ_0	1.306	0.118	0.172	0.159
	(1.121)	(0.004)	(0.038)	(0.026)
γ_{11}	–	0.002	–	–
	–	(2.58e-4)	–	–
γ_{12}	0.001	0.003	–	–
	(5.64e-4)	(2.24e-4)	–	–
γ_{13}	0.006	0.006	–	–
	(1.18e-4)	(1.42e-4)	–	–
$\lambda^1_{Y,11}$	–	0.629	–	-0.018
	–	(0.290)	–	(0.020)
$\lambda^1_{Y,22}$	–	0.150	–	-0.703
	–	(0.101)	–	(0.310)
σ_y^2	8.34e-8	8.01e-8	8.02e-8	7.84e-8
	(1.42e-9)	(1.33e-9)	(1.44e-9)	(1.40e-9)
Loglik y_t^τ	35443	35530	35315	35341
Loglik v_t^τ	–	–	26264	26260

Table 4 In-sample model fit

This table gives the in-sample fit for four model specifications presented in Table 3. The performance of our model (G_3SV_3) in matching yields is compared with the purely Gaussian setting (G_3SV_0). Additionally, we report its ability to match the volatility dynamics. In panels $a1$ and $b1$, we report root mean squared error (RMSE) in bps per annum. The fit of volatilities is also stated as RMSE. To obtain easily comparable numbers, we convert the covariance matrix of yields into covolatilities: Since the slope and covariance can be negative, we take the square root of their absolute values. This serves as an input for computing the RMSE for the term structure of volatilities. Panels $a2$ and $b2$ show the percentage of variation in observed yields and volatilities explained by the respective model-implied quantities. The numbers represent the R^2's from the regression of the observed variable on the fitted one.

	G_3SV_0 risk-neutral	G_3SV_0	G_3SV_3 risk-neutral	G_3SV_3
Panel a. Term structure of yields				
$a1$. RMSE in bps				
6 months	1.3	1.3	1.3	1.5
2 year	2.4	2.3	2.3	2.2
3 year	1.6	1.6	1.5	1.6
5 year	2.9	2.8	2.7	2.6
7 year	1.8	1.7	1.7	1.7
10 year	2.8	2.8	2.5	2.5
$a2$. Explained variation (%)				
6 months	99.99	99.99	99.99	99.99
2 year	99.97	99.97	99.98	99.98
3 year	99.99	99.99	99.99	99.99
5 year	99.94	99.94	99.95	99.97
7 year	99.97	99.98	99.98	99.99
10 year	99.93	99.93	99.94	99.96
Panel b. Term structure of volatilities				
$b1$. RMSE in bps				
v_t^{2Y}	–	–	6.1	5.8
$v_t^{10Y} - v_t^{2Y}$	–	–	6.1	5.8
$v_t^{5Y,10Y}$	–	–	4.7	4.7
$b2$. Explained variation (%)				
v_t^{2Y}	–	–	95.92	96.22
$v_t^{10Y} - v_t^{2Y}$	–	–	98.23	98.30
$v_t^{5Y,10Y}$	–	–	96.92	96.89

Table 5 Out-of-sample yield forecasting

This table reports RMSEs in bps for forecasting horizons $h = 1, 4, 12$ and 52 weeks. The models are estimated using weekly data from 1992:01 through 2004:12. Forecasts are performed over the period 2005:01 through 2007:12.

Panel *a*. Yield forecasting

Yield Maturity	Forecast Horizon (h weeks)	Random Walk	G_3SV_0	G_3SV_3
6 months	1	9.29	13.54	13.62
2 years	1	10.93	14.89	15.10
3 years	1	11.36	14.80	14.81
5 years	1	10.77	14.13	14.30
7 years	1	9.98	13.56	13.67
10 years	1	9.17	13.11	13.59
6 months	4	21.16	23.03	22.81
2 years	4	22.74	26.99	27.16
3 years	4	23.27	27.04	27.08
5 years	4	22.37	25.63	25.85
7 years	4	20.88	24.11	24.38
10 years	4	19.08	22.27	23.23
6 months	12	43.18	39.69	38.66
2 years	12	43.47	43.16	43.33
3 years	12	42.40	42.37	42.58
5 years	12	38.28	38.99	39.80
7 years	12	34.64	36.12	37.53
10 years	12	31.05	33.19	36.21
6 months	52	146.56	89.37	82.77
2 years	52	110.02	70.27	67.87
3 years	52	92.73	63.88	63.57
5 years	52	66.82	53.43	59.00
7 years	52	53.12	47.10	57.99
10 years	52	44.64	41.74	60.48

Panel *b*. Volatility forecasting

Volatility Maturity	Forecast Horizon (h weeks)	Random Walk	G_3SV_0	G_3SV_3
2 years	1	11.28	n.a.	13.62
5 years	1	11.75	n.a.	10.96
10 years	1	9.76	n.a.	9.71
2 years	4	25.81	n.a	22.33
5 years	4	26.12	n.a.	16.47
10 years	4	22.78	n.a.	15.19
2 years	12	33.14	n.a.	29.08
5 years	12	34.13	n.a.	26.35
10 years	12	30.90	n.a.	26.41

Table 6 Regressions of yield realized second moments on filtered states

This table reports the t-statistics and adjusted R^2's obtained by regressing the yield realized variances (first five columns) and covariances (last three columns) on the latent factors extracted from our model, i.e. $X_1, X_2, f, V_{11}, V_{12}, V_{22}$. The sample covers the 1992:01–2007:12 period, which amounts to 845 weekly observations. t-statistics are corrected using the Newey-West adjustment with 12 lags.

RV & RCov	2Y	3Y	5Y	7Y	10Y	(5Y,2Y)	(10Y,2Y)	(10Y,5Y)
const.	2.53	2.86	-0.08	1.45	1.32	2.52	-0.10	-0.03
X_1	1.27	-2.56	1.03	-1.17	1.44	0.52	0.73	-2.51
X_2	1.25	1.94	1.12	2.45	1.16	1.32	-2.22	0.17
f	3.83	1.25	0.68	0.61	3.75	2.75	-0.13	-2.74
V_{11}	5.81	3.70	7.74	9.06	19.48	13.76	12.90	17.46
V_{12}	28.52	8.58	19.50	24.96	29.54	34.95	27.87	68.45
V_{22}	43.01	29.89	12.67	11.74	7.43	25.12	16.29	25.78
Adj. R^2	0.97	0.88	0.87	0.90	0.93	0.94	0.93	0.97

Table 7 Regressions of filtered states on yield curve principal components

This table reports regressions of model-implied factors on principal components extracted from the yield curve. The sample covers the 1992:01–2007:12 period, which amounts to 845 weekly observations. Standard errors are corrected using Newey-West adjustment with 12 lags and are reported in parentheses. Both left- and right-hand side variables are standardized to make the regression coefficients directly comparable. X_1 and X_2 have been multiplied by -1.

Regressors	X_1	X_2	f	V_{11}	V_{12}	V_{22}
PC_1	0.544	0.784	0.875	-0.102	0.013	-0.138
	(0.003)	(0.004)	(0.001)	(0.070)	(0.068)	(0.087)
PC_2	0.818	-0.124	-0.458	0.533	-0.397	0.358
	(0.004)	(0.005)	(0.001)	(0.080)	(0.085)	(0.079)
PC_3	0.184	-0.604	0.155	0.158	-0.233	0.184
	(0.003)	(0.005)	(0.001)	(0.061)	(0.057)	(0.086)
PC_4	-0.002	0.028	-0.012	-0.345	0.463	-0.205
	(0.003)	(0.004)	(0.001)	(0.062)	(0.064)	(0.082)
PC_5	-0.012	0.013	0.006	-0.182	0.151	-0.109
	(0.003)	(0.005)	(0.001)	(0.060)	(0.057)	(0.089)
Adj. R^2	1.00	1.00	1.00	0.47	0.45	0.23

Table 8 Regressions of filtered states on interest rate forecasts

This table reports regressions of the model-implied factors on survey-based expectations and uncertainty about US interest rates. Monthly factors are obtained by averaging the weekly numbers returned from the estimation. As explanatory variables, we use yield curve forecasts compiled by the BCFF survey. $E(\cdot)$ denotes the consensus, defined as a median forecast; $\sigma(\cdot)$ proxies for the uncertainty, and is computed as the mean absolute deviation of individual forecasts. The sample covers the 1992:01–2007:12 period, which amounts to 192 monthly observations. Standard errors are corrected using Newey-West adjustment with four lags and are reported in parentheses. Both the latent factors and the survey data are standardized in order to make the regression coefficients directly comparable. The qualifiers $hQ1, hQ4$ at each coefficient denote the horizon for which panelists in BCFF report their forecast. $hQ1$ is the one-quarter-ahead prediction (i.e. prediction for the next quarter), $hQ4$ is the four-quarter-ahead prediction. Table contains results only for significant variables. For the ease of interpretation, X_1 and X_2 have been multiplied by -1, so that they have a positive correlation with the ten-year yield and two-year yield, respectively.

Regressors	X_1	X_2	f	V_{11}	V_{12}	V_{22}
E(FFR)	–	–	0.967^{hQ1}	–	–	–
	–	–	(0.066)	–	–	–
E(y2Y)	–	1.558^{hQ1}	–	–	–	–
	–	(0.172)	–	–	–	–
σ(y2Y)	–	–	–	–	–	0.430^{hQ1}
	–	–	–	–	–	(0.101)
E(y5Y)	–	0.763^{hQ1}	–	–	–	–
	–	(0.162)	–	–	–	–
σ(y5Y)	–	–	–	1.040^{hQ4}	-0.921^{hQ1}	–
	–	–	–	(0.244)	(0.170)	–
E(y10Y)	1.035^{hQ4}	–	–	–	–	–
	(0.093)	–	–	–	–	–
σ(y10Y)	–	–	–	-0.483^{hQ4}	0.325^{hQ1}	–
	–	–	–	(0.129)	(0.111)	–
E(y30Y)	-0.216^{hQ4}	–	–	–	–	–
	(0.092)	–	–	–	–	–
σ(y30Y)	–	–	–	-0.520^{hQ4}	0.634^{hQ1}	–
	–	–	–	(0.251)	(0.136)	–
Adj. R^2	0.76	0.78	0.93	0.20	0.22	0.18

Table 9 Regressions of filtered states on macroeconomic surveys

This table reports regressions of the model-implied factors on macroeconomic surveys. Monthly factors are obtained by averaging the weekly numbers returned from the estimation. We splice monthly data from two surveys: BCFF and BCEI compiled over the period 1992:01–2007:12. This amounts to 192 monthly observations. The following variables are used: real GDP (RGDP), unemployment (UNEMPL), housing starts (HOUST), federal funds rate (FFR), industrial production (IP), consumer price index (CPI). RGDP and FFR forecasts are obtained from BCFF, while the remaining variables are from BCEI. $E(\cdot)$ denotes the consensus, defined as a median forecast; $\sigma(\cdot)$ proxies for the uncertainty, and is computed as the mean absolute deviation of individual forecasts. Standard errors are corrected using Newey-West adjustment with four lags and are reported in parentheses. Both the latent factors and the survey data are standardized in order to make the regression coefficients directly comparable. The table contains only significant variables. For the ease of interpretation, X_1 and X_2 have been multiplied by -1, so that they have a positive correlation with the two-year yield and ten-year yield, respectively.

Regressor	X_1	X_2	f	V_{11}	V_{12}	V_{22}
E(RGDP)	0.350	0.236	–	0.664	–	–
	(0.065)	(0.077)	–	(0.065)	–	–
σ(RGDP)	0.248	0.188	–	–	-0.257	–
	(0.063)	(0.057)	–	–	(0.117)	–
E(UNEMPL)	0.770	–	–	0.238	–	–
	(0.093)	–	–	(0.089)	–	–
σ(UNEMPL)	–	–	–	0.385	-0.197	–
	–	–	–	(0.081)	(0.100)	–
E(HOUST)	-0.417	–	–	-0.310	–	-0.315
	(0.098)	–	–	(0.080)	–	(0.052)
σ(HOUST)	–	-0.232	–	–	-0.386	–
	–	(0.053)	–	–	(0.079)	–
E(FFR)	0.516	0.862	0.868	–	–	-0.285
	(0.086)	(0.088)	(0.026)	–	–	(0.098)
σ(FFR)	–	–	0.041	–	-0.144	0.299
	–	–	(0.021)	–	(0.078)	(0.073)
E(IP)	–	–	-0.108	–	–	–
	–	–	(0.028)	–	–	–
σ(IP)	–	–	-0.057	–	–	0.305
	–	–	(0.023)	–	–	(0.078)
E(CPI)	–	–	0.074	–	–	–
	–	–	(0.021)	–	–	–
σ(CPI)	–	-0.172	–	–	–	0.254
	–	(0.061)	–	–	–	(0.071)
Adj. R^2	0.81	0.73	0.96	0.52	0.29	0.47

References

Adrian, T., Wu, H., 2009. The term structure of inflation expectations. Unpublished working paper. Federal Reserve Bank of New York.

Andersen, T. G., Benzoni, L., 2009. Do bonds span volatility risk in the US Treasury market? A specification test for affine term structure models. Journal of Finance, forthcoming.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. Journal of International Economics 73, 251–277.

Ang, A., Boivin, J., Dong, S., Loo-Kung, R., 2009. Monetary policy shifts and the term structure. Unpublished working paper. Columbia University, HEC Montréal, Ortus Capital Management.

Ang, A., Dong, S., Piazzesi, M., 2007. No-arbitrage Taylor rules. Unpublished working paper. Columbia University, University of Chicago, NBER and CEPR.

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure with macroeconomic and latent variables. Journal of Monetary Economics 50, 745–787.

Audrino, F., Corsi, F., 2007. Realized correlation tick-by-tick. Unpublished working paper. University of St. Gallen.

Audrino, F., Corsi, F., 2008. Realized covariance tick-by-tick in presence of rounded time stamps and general microstructure effects. Discussion paper No. 2008-04, University of St. Gallen.

Bandi, F., Russell, J., 2008. Microstructure noise, realized variance, and optimal sampling. Review of Economic Studies 75, 339–369.

Barndorff-Nielsen, O. E., Shephard, N., 2004. Econometric analysis of realized covariation: High-frequency based covariance, regression and correlation in financial economics. Econometrica 72, 885–925.

Beechey, M. J., Wright, J. H., 2008. The high-frequency impact of news on long-term yields and forward rates: Is it real? Unpublished working paper. Sveriges Riksbank and Johns Hopkins University.

Bekker, P. A., Bouwman, K. E., 2009. Risk-free interest rates driven by capital market returns. Unpublished working paper. University of Groningen and Erasmus University Rotterdam.

Bikbov, R., Chernov, M., 2008. No-arbitrage macroeconomic determinants of the yield curve. Unpublished working paper. London Business School and CEPR.

Bikbov, R., Chernov, M., 2009. Unspanned stochastic volatility in affine models: Evidence from Eurodollar futures and options. Management Science, forthcoming.

Bru, M.-F., 1991. Wishart processes. Journal of Theoretical Probability 4, 725–751.

Buraschi, A., Cieslak, A., Trojani, F., 2008. Correlation risk and the term structure of interest rates. Unpublished working paper. University of Lugano.

Buraschi, A., Porchia, P., Trojani, F., 2009. Correlation risk and optimal portfolio choice. Journal of Finance, forthcoming.

Campbell, J. Y., Shiller, R. J., 1991. Yield spreads and interest rate movements: A bird's eye view. Review of Economic Studies 58, 495–514.

Campbell, J. Y., Sunderam, A., Viceira, L. M., 2009. Inflation bets or deflation hedges? The changing risk of nominal bonds. Unpublished working paper. Harvard Business School.

Carr, P., Wu, L., 2007. Stochastic skew in currency options. Journal of Financial Economics 86, 213–247.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B., 1992. An empirical comparison of alternative models of the short-term interest rate. Journal of Finance 47, 1209–1227.

Christoffersen, P., Jacobs, K., Karoui, L., Mimouni, K., 2009. Non-linear filtering in affine term structure models: Evidence from the term structure of swap rates. Unpublished working paper. McGill University, Goldman Sachs, American University in Dubai.

Cochrane, J. H., Piazzesi, M., 2005. Bond risk premia. American Economic Review 95, 138–160.

Cochrane, J. H., Piazzesi, M., 2008. Decomposing the yield curve. Unpublished working paper. University of Chicago.

Collin-Dufresne, P., Goldstein, R. S., 2002. Do bonds span the fixed income markets? Theory and evidence for the unspanned stochastic volatility. Journal of Finance 58, 1685–1730.

Collin-Dufresne, P., Goldstein, R. S., Jones, C. S., 2008. Can interest rate volatility be extracted from the cross-section of bond yields? Journal of Financial Economics, forthcoming.

da Fonseca, J., Grasselli, M., Tebaldi, C., 2006. Option pricing when correlations are stochastic: An analytical framework. Unpublished working paper. ESLIV, University of Padova and University of Verona.

Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

Drost, F. C., Nijman, T. E., 1993. Temporal aggregation of GARCH processes. Econometrica 61, 909–927.

Duffee, G. R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.

Duffee, G. R., 2006. Term structure estimation without using latent factors. Journal of Financial Economics 79, 507–536.

Duffee, G. R., 2007. Are variations in term premia related to the macroeconomy? Unpublished working paper. University of California, Berkeley.

Dungey, M., McKenzie, M., Smith, V., 2009. Empirical evidence on jumps in the term structure of the US Treasury market. Journal of Empirical Finance 16, 430–445.

Engle, R. F., Sheppard, K., 2001. Theoretical and emprical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper No. 8554.

Fengler, M. R., Härdle, W. K., Villa, C., 2001. The dynamics of implied volatilities: A common principal components approach. Unpublished working paper. Humboldt-Universität zu Berlin, University of Rennes.

Fisher, M., Gilles, C., 1996. Estimating exponential-affine models of the term structure. Unpublished working paper. Federal Reserve Bank of Atlanta.

Fisher, M., Nychka, D., Zervos, D., 1994. Fitting the term structure of interest rates with smoothing splines. Unpublished working paper. Federal Reserve and North Carolina State University.

Fleming, M. J., 1997. The round-the-clock market for US Treasury securities. FRBNY Economic Policy Review.

Fleming, M. J., 2003. Measuring Treasury market liquidity. FRBNY Economic Policy Review.

Fleming, M. J., Mizrach, B., 2008. The microstructure of a US Treasury ECN: The BrokerTec platform. Unpublished working paper. Federal Reserve Bank of New York and Rutgers University.

Gourieroux, C., 2006. Continuous time Wishart process for stochastic risk. Econometric Reviews 25, 177–217.

Gourieroux, C., Jasiak, J., Sufana, R., 2005. The Wishart autoregressive process of multivariate stochastic volatility. Unpublished working paper. CREST, CEPREMAT, and University of Toronto.

Greenspan, A., 2009. The Fed didn't cause the housing bubble. Wall Street Journal March 11, Eastern Edition, A15.

Gurkaynak, R. S., Sack, B., Wright, J. H., 2006. The US Treasury yield curve: 1961 to the present. Unpublished working paper. FRB Washington.

Han, B., 2007. Stochastic volatilities and correlations of bond yields. Journal of Finance 62, 1491–1524.

Haubrich, J., Pennacchi, G., Ritchken, P., 2008. Estimating real and nominal term structures using Treasury yields, inflation, inflation forecasts, and inflation swap rates. Unpublished working paper. Federal Reserve Bank of Cleveland.

Hautsch, N., Ou, Y., 2008. Yield curve factors, term structure volatility, and bond risk premia. Unpublished working paper. Humboldt University Berlin.

Hayashi, T., Yoshida, N., 2005. On covariance estimation of non-synchronously observed diffusion processes. Bernoulli 11, 359–379.

Heidari, M., Wu, L., 2003. Are interest rate derivatives spanned by the term structure of interest rates? Journal of Fixed Income 13, 75–86.

Jacobs, K., Karoui, L., 2009. Conditional volatility in affine term-structure models: Evidence from Treasury and swap markets. Journal of Financial Economics 91, 288–318.

Jacod, J., 1994. Limit of random measures associated with the increments of a Brownian semimartingale. Unpublished working paper. Universite Pierre et Marie Curie.

Jiang, G. J., Lo, I., Verdelhan, A., 2009. Information shocks, jumps, and price discovery: Evidence from the US Treasury market. Journal of Financial and Quantitative Analysis, forthcoming.

Jones, C. M., Lamont, O., Lumsdaine, R. L., 1998. Macroeconomic news and bond market volatility. Journal of Financial Economics 47, 315–337.

Joslin, S., 2006. Can unspanned stochastic volatility models explain the cross section of bond volatilities? Unpublished working paper. MIT Sloan School of Management.

Joslin, S., 2007. Pricing and hedging volatility risk in fixed income markets. Unpublished working paper. MIT Sloan School of Management.

Julier, S., Uhlmann, J., 1997. A new extension of the Kalman filter to nonlinear systems. In Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls.

Kalman, R., 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82, 35–45.

Kim, D. H., 2007a. Challenges in macro-finance modelling. BIS Working Paper, No. 240.

Kim, D. H., 2007b. Spanned stochastic volatility in bond markets: A re-examination of the relative pricing between bonds and bond options. BIS Working Paper Series No. 239.

Li, H., Zhao, F., 2006. Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. Journal of Finance 61, 341–378.

Litterman, R., Scheinkman, J. A., 1991. Common factors affecting bond returns. Journal of Fixed Income 1, 54–61.

Magnus, J. R., Neudecker, H., 1979. The commutation matrix: Some properties and applications. Annals of Statistics 7, 381–394.

Mishkin, F. S., 2007. Housing and the monetary transmission mechanism. Finance and Economics Discussion Series, Federal Reserve Board.

Mizrach, B., Neely, C. J., 2006. The transition to electronic communications networks in the secondary Treasury market. Federal Reserve Bank of St. Louis Review.

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

Phoa, W., 1997. Can you derive market volatility forecast from the observed yield curve convexity bias? Journal of Fixed Income, 43–53.

Price, K. V., Storn, R. M., Lampinen, J. A., 2005. Differential Evolution: A Practical Approach to Global Optimization. Springer Berlin.

Thompson, S., 2008. Identifying term structure volatility from the LIBOR-swap curve. Review of Financial Studies 21, 819–854.

Trolle, A. B., Schwartz, E. S., July 2009. A general stochastic volatility model for the pricing of interest rate derivatives. Review of Financial Studies 22, 2007–2057.

Van Loan, C. F., 1978. Computing integrals involving matrix exponential. IEEE Transactions on Automatic Control 23, 395–404.

Waggoner, D. F., 1997. Spline methods for extracting interest rate curves from coupon bond prices. Unpublished working paper. FRB of Atlanta.

Wan, E. A., van der Merwe, R., 2001. Kalman Filtering and Neural Networks. John Wiley & Sons, Inc.

Wright, J., Zhou, H., 2009. Bond risk premia and realized jump risk. Journal of Banking and Finance, forthcoming.

Zhang, L., Mykland, P. A., Ait-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.