Modeling the term structure of interest rates

DISSERTATION
submitted by
PAVOL POVALA
for the degree of
Ph.D. in Economics
at the Faculty of Economics
University of Lugano (USI)
Switzerland

THESIS COMMITTEE:
Prof. Patrick Gagliardini, University of Lugano (SFI)
Prof. Stijn Van Nieuwerburgh, New York University, Stern School of Business
Prof. Paul Schneider, University of Lugano (SFI)
Prof. Fabio Trojani, University of Lugano (SFI)

September 27, 2013

For Sophia, love of my life

Acknowledgement

I would like to express my gratitude to people who contributed significantly to this thesis: Anna Cieslak, Stijn Van Nieuwerburgh, Fabio Trojani, and Tomasz Wisniewski.

Contents

List of Tables

List of Figures

Introduction

This doctoral thesis studies the behavior of Treasury yields. I focus on three aspects of the yield curve: (*i*) the link between bond yields and key macro quantities such as consumption growth, (*ii*) the process of expectations formation about the short term interest rate which, in most developed markets, is the main monetary policy instrument, and (*iii*) the behavior of interest rate volatilities across different maturities. These research questions share a common goal which is to provide guidance for designing macroeconomic models of the yield curve.

The economic theory has clear predictions about the joint behavior of macro quantities and the yield curve. However, it has been hard to reconcile these predictions with the observed dynamics of yields. Chapter 1: "Intertemporal Trade-off and the Yield Curve" empirically evaluates and models the relationship between consumption growth and the short term riskless interest rate, charaterized by the elasticity of intertemporal substitution (EIS). Chapter 2: "Expecting the Fed" studies how investors form their expectations about the short rate. In particular, it evaluates if and how investors' expectations deviate from the benchmark of full information rational expectations. Chapter 3: "Information in the term structure of yield curve volatility" analyzes the dynamic features of yield volatilities and undertakes their modeling. Chapters 2 and 3 are based on joint work with Anna Cieslak.

Understanding the relationship between the short term real rate and the consumption/output growth is important for several reasons. First, the EIS is one of the most important parameters in asset pricing and macroeconomics. In consumption-based models, the EIS is a key determinant of the level and variation in the risk-free rate. Second, it determines the magnitude of various policy effects such as the real effects of monetary policy. Despite the importance of this relationship, the literature has not reached consensus on the value of the EIS. On one hand, estimates from aggregate data suggest that the EIS is close to zero (Hall, 1988; Campbell, 2003; Fuhrer and Rudebusch, 2004). On the other hand, evidence from disaggregated data indicates that the EIS is likely to be above unity for households with substantial asset holdings (Vissing-Jorgensen and Attanasio, 2003; Gruber, 2006). Micro evidence is often used to motivate the EIS above unity in calibrations of asset pricing models (Barro, 2009; Bansal, Kiku, and Yaron, 2012).

Chapter 1 provides two novel empirical results that help resolve the conflicting and inconclusive evidence on the magnitude of the EIS.

First, I propose a new measure of the ex-ante real interest rate that relies solely on asset prices and avoids using realized inflation. Specifically, I show that the difference between the three-month and ten-year nominal Treasury yield, which is the negative of the yield curve slope, closely tracks the variation in the ex-ante short-term real rate. This measure has two properties that are useful for the estimation of the EIS. It does not require any measure of expected inflation thus avoiding short samples with direct measures of inflation expectations or auxiliary assumptions on the time series dynamics of inflation such as linearity or rational expectations.

Second, EIS estimates using aggregate consumption data are consistently negative and statistically significant around -0.5. Similarly, replacing aggregate consumption by the aggregate output data in the Euler equation yields negative estimates with magnitudes between -0.5 and -1.2.

I show that the negatively signed estimates of the EIS arise naturally in a limited market participation model. The key mechanism that changes the sign of the EIS operates through the interaction of two types of households. A fraction of households, rule-of-thumb consumers, do not participate in asset markets and consume their wage income every period (Campbell and Mankiw, 1989; Mankiw, 2000). The remaining households, savers, have positive EIS and own all the productive assets in the economy. Savers have two sources of income, wages and dividends. Changes in the real interest rate directly influence only the consumption and labor supply of savers. But the intertemporal choices of savers then have an impact on the equilibrium real wage of rule-of-thumb consumers and thereby alter their consumption. Changes in the real wage also lead to variation in profit margins of firms, affecting the dividend income and the consumption of savers. The resulting effect of the real interest rate shock on aggregate consumption is negative if the fraction of rule-of-thumb households is sufficiently high and the labor supply is sufficiently inelastic.

The negative relationship between consumption growth and ex-ante real interest rates has a number of important implications for both asset pricing and monetary policy.

First, the negative value of the aggregate EIS helps explain why the yield curve behavior is hard to reconcile with asset pricing models (Duffee, 2012a). Consumption-based models calibrated with a positive EIS imply a short-term real interest rate that is negatively correlated with the real interest rate from the data (Cumby and Diba, 2007). Intuitively, changing the sign of the EIS aligns the model-implied real rate with the ex-ante real rate from the data.

Second, persistent but mean-reverting variation in the real interest rate together with sizeable estimates of the EIS imply a time-varying expected consumption growth rate. Using this link, I extract the persistent component of aggregate consumption growth and study its asset pricing implications. In line with the intuition of long-run risk models, the extracted component has a high persistence, explains about 11% of the variation in consumption growth and is positively related to the price-dividend ratio. In contrast to the calibration favored by the long-run risk literature, the persistent component that is consistent with asset prices is obtained with the negative value of the EIS.

The negative EIS also matters for models of monetary policy because the central bank reacts to aggregate output. The aggregate Euler equation (IS curve) with a positive EIS is the core of dynamic general equilibrium models used in the monetary policy analysis, e.g. New-Keynesian models. The sign and magnitude of the aggregate EIS determine the real effect of monetary policy. Therefore, the negative sign of the EIS changes key predictions of these models. Most notably, it implies that an increase in the real interest rate is expansionary.

Chapter 2 focuses on how market participants form and update their expectations about the short term interest rate. To directly disentangle the risk premium from short rate expectations, we rely on comprehensive survey data containing the term structure of private sector's forecasts of the federal funds rate (FFR) as well as forecasts of longer maturity yields. Separating short rate expectations from the risk premia in Treasury bonds is important for policy makers and for understanding the economics of the yield curve. Such decomposition provides insight about how markets perceive the future course of monetary policy, economic activity, inflation and their associated risks. The results suggest that the view of frictionless rational expectations deviates from the observed behavior of interest rates in significant ways.

First, while survey-based short rate expectations match almost one-to-one the contemporaneous behavior of short-term yields and fed fund futures, these expectations are poor predictors of future short rates except at very short horizons (e.g. Rudebusch, 2002). Agents, faced with a highly persistent process, fail in real time to appraise the mean reversion in the policy rate that occurs at the business cycle frequency. As such, agents' ex-post forecast errors of the short rate are predictable with past information that captures the mean reversion component. Our evidence suggests that this feature of short rate expectations pertains to an environment with an active central bank that itself might adapt its policy rule over time, and is less likely to characterize the data pre-Fed. In the last three decades, we find that agents' forecast errors about the short rate comove closely (with a negative sign) with the errors they make when forecasting unemployment, and much less so inflation.

Second, we construct a measure of expectations frictions based on the difference between the observed physical dynamics of the policy rate and the corresponding expectations reflected in the yield curve. This variable reflects the idea that there is information in the time series of monetary policy actions that is not fully impounded in the cross section of yields in real time. We label this variable as $MP_t^{\perp}$. One interpretation is that the Fed is able to deliver persistent surprises to the market. Accordingly, we show that $MP_t^{\perp}$ picks up the low-frequency movements in the measure of monetary policy surprises identified from high frequency data (Kuttner, 2001), which the literature has documented to be unaffected by the risk premium (Piazzesi and Swanson, 2008).

With the help of survey data for longer-maturity yields, we obtain a model-free decomposition of excess bond returns into a risk premium and an ex-ante unexpected return component. The unexpected return on a two-year bond moves in lockstep with the (negative of) FFR forecast errors, and one quarter of its variation can be predicted ex-post by $MP_t^{\perp}$. The effect of expectations frictions is the strongest at the short end of the maturity spectrum, most influenced by monetary policy, and subsides for long-term bonds. We find that several conditioning variables used to forecast realized bond returns, especially variables related to the real activity, predict unexpected returns and are essentially uncorrelated with the survey-implied risk premia. These variables feature a high degree of correlation with $MP_t^{\perp}$.

Chapter 3 studies the structure, economic content and pricing implications of the fluctuating covariance matrix of interest rates. Using comprehensive data on transactions in the US Treasury market, we explore the key drivers of volatility in government bonds. To this end, we decompose the stochastic covariance matrix of on-the-run yields into two volatility components corresponding to yield factors with a short and long duration, respectively, and an additional variable that captures the degree of comovement between the short and long maturity segment of the yield curve. This decomposition delivers the following results.

First, we find a lead-lag relationship between yield volatility and market-wide liquidity. The long-duration volatility component captures almost half of the variation in a broad measure of the value of funding liquidity. That element of liquidity is visible in the on-/off-the-run premium across bond maturities and reflects the tightness of financial conditions in asset markets (Fontaine and Garcia, 2011). Typically, it also preempts an action of the monetary authority. An increase in the long-duration volatility component predicts a decline in the value of funding liquidity up to six months ahead with an R^2 of nearly 50%. The short-duration volatility, instead, dominates in explaining the transitory episodes of liquidity dry-ups that do not necessarily trigger a monetary policy reaction (e.g. GM/Ford downgrade), but reflect the amount of arbitrage capital available in the market (Hu, Pan, and Wang, 2013). Importantly, an increase in the short-end volatility component predicts that the arbitrage capital will be scarce in the future.

Second, we explore the macroeconomic drivers of volatility in yields. A combination of expectations and uncertainty about monetary policy, inflation and real activity accounts for up to 30% of the variation in yield volatility. Short- and long-run volatility are related to different macro variables. The long-end volatility shows a pronounced response to the real activity measures such as the uncertainty about unemployment, and to expectations about the path of monetary policy. Quite differently, the short-term volatility is most strongly linked to the monetary policy uncertainty. Finally, the comovement term is associated with the uncertainty about inflation, monetary policy and the real economy, showing that those variables influence the extent of correlation between yields across maturities.

Third, we study how interest rate volatility is related to the risk compensation required by investors for holding Treasury bonds. When short-duration volatility is high, as induced by an increased monetary policy uncertainty, bond risk premia tend to rise. This effect is statistically and economically strong for bonds with a two-year maturity, but it dissipates quickly as the maturity increases.

The chapters follow the inverse chronological order. Appendices are collected at the end of each chapter.

Chapter 1

Intertemporal Trade-off and the Yield Curve

I. Introduction

Intertemporal trade-off is one of the key concepts in asset pricing and macroeconomics.[1] The trade-off is described by the elasticity of intertemporal substitution (EIS) defined as the change in expected consumption growth in response to a change in the real interest rate. The EIS does not only have a large impact on the behavior of asset prices but it also determines the magnitude of various policy effects. In consumption-based models, the willingness of agents to shift consumption intertemporally is a key determinant of the level and variation in yield on the risk-free asset.

Despite its importance, the literature has not yet settled on the value of the EIS. On one hand, estimates from aggregate data suggest that the EIS is close to zero (Hall, 1988; Campbell, 2003; Fuhrer and Rudebusch, 2004). On the other hand, evidence from disaggregated data indicates that the EIS is likely to be above unity for households with substantial asset

[1]This chapter is based on the research paper with the same title. This chapter was written while I was a visiting scholar at the Finance Department of Stern School of Business, NYU whose hospitality is gratefully acknowledged. I received helpful comments from Luca Benzoni, Itamar Drechsler, Greg Duffee, Douglas Gale, Ralph Koijen, Anthony Lynch, Matteo Maggiori, Alexi Savov, Fabio Trojani, Stephen Wright, Liuren Wu, participants of the Financial Economics Workshop at NYU, seminar participants at Birkbeck, Chicago Fed, Fed Board, Stockholm School of Economics, and Bank of Canada. I am especially grateful to Anna Cieslak and Stijn Van Nieuwerburgh for numerous comments and discussions. This research project has been carried out within the project on "New Methods in Theoretical and Empirical Asset Pricing" of the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK). The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

holdings (Vissing-Jorgensen and Attanasio, 2003; Gruber, 2006). Micro evidence is often used to motivate the EIS above unity in calibrations of asset pricing models (Barro, 2009; Bansal, Kiku, and Yaron, 2012).

This paper provides two novel empirical results that help resolve the conflicting and inconclusive evidence on the magnitude of the EIS.

First, I propose a new concept for measuring the ex-ante real interest rate that relies solely on asset prices and avoids using realized inflation. Specifically, I show that the difference between the three-month and ten-year nominal Treasury yield, which is the negative of the yield curve slope, closely tracks the variation in the ex-ante short-term real rate. I label this new measure of the ex-ante real rate as $\tilde{r}_t = -\text{slope}_t$. This measure has two properties that are useful for the estimation of the EIS. It does not require any measure of expected inflation thus avoiding short samples with direct measures of inflation expectations or auxiliary assumptions on the time series dynamics of inflation such as linearity or rational expectations. As a result, my measure is available for a long sample and has favorable stationarity properties across different monetary policy regimes. Given that the nominal Treasury yield curve data are available for long periods, using $\tilde{r}_t$ allows me to estimate the EIS from the Euler equation using 135 years of data.

Second, EIS estimates using aggregate consumption data are consistently negative and statistically significant around -0.5. Similarly, replacing aggregate consumption by the aggregate output data in the Euler equation yields negative estimates with magnitudes between -0.5 and -1.2. The results hold across sub-samples marked by different monetary policies and institutional arrangements. Hence, the stability of these EIS estimates requires a structural interpretation of the negative sign.

When is the yield curve slope a good measure of the real rate variation? Three restrictions on statistical properties of inflation expectations, risk premia and the real rate need to hold so that the slope of the nominal yield curve closely tracks the ex-ante real rate, all of which are met with theoretical and empirical support. First, the Fisher effect holds in the long run, i.e. the real interest rate is stationary and most of the variation in the long-term nominal yield is due to changing inflation expectations. Second, inflation expectations can be described by a single factor with a close-to-unit-root persistence. Third, the variation in risk premia is small and less persistent than the variation of the real rate.

I show that the negative sign of the EIS arises naturally in a standard limited market participation model. The key mechanism that changes the sign of the EIS operates through the interaction of two types of households. A fraction of households, rule-of-thumb consumers, do not participate in asset markets and consume their wage income every period (Campbell and Mankiw, 1989; Mankiw, 2000). The remaining households, savers, have positive EIS and own all the productive assets in the economy. Savers have two sources of income, wages and dividends. Changes in the real interest rate directly influence only the consumption and labor supply of savers. But the intertemporal choices of savers then have an impact on the equilibrium real wage of rule-of-thumb consumers and thereby alter their consumption. Changes in the real wage also lead to the variation in profit margins of firms, affecting the dividend income and the consumption of savers. The resulting effect of the real interest rate shock on aggregate consumption is negative if the fraction of rule-of-thumb households is sufficiently high and the labor supply is sufficiently inelastic. Standard values for the structural parameters of the model that are consistent with micro evidence imply the estimated value of the aggregate EIS.

The negative relationship between consumption growth and ex-ante real interest rates has a number of important implications for both asset pricing and monetary policy.

First, the negative value of the aggregate EIS helps explain why the yield curve behavior is hard to reconcile with asset pricing models (Duffee, 2012a). Consumption-based models calibrated with a positive EIS imply a short-term real interest rate that is negatively correlated with the real interest rate from the data (Cumby and Diba, 2007).[2] Intuitively, changing the sign of the EIS aligns the model-implied real rate with the ex-ante real rate from the data.

Second, persistent but mean-reverting variation in the real interest rate together with sizeable estimates of the EIS imply a time-varying expected consumption growth rate. Using this link, I extract the persistent component of aggregate consumption growth and study its asset pricing implications. In line with the intuition of long-run risk models, the extracted component has a high persistence, explains about 11% of the variation in consumption growth and is positively related to the price-dividend ratio. New and in contrast to the calibration favored by the long-run risk literature, the persistent component that is consistent with asset prices is obtained with the negative value of the EIS.

[2] Cumby and Diba (2007) compare the model-implied real rate to the ex-ante and ex-post real rate obtaining similar results for both.

The implications for consumption-based asset pricing can be summarized as follows. To replicate the dynamics of asset prices one can either use the aggregate consumption data together with a negative value of the EIS in the representative agent Epstein-Zin framework, or one can use a positive value of the EIS close to unity and apply it only to the consumption of households participating in asset markets along the lines of Malloy, Moskowitz, and Vissing-Jorgensen (2009).

The negative EIS also matters for models of monetary policy because the central bank reacts to aggregate output. The aggregate Euler equation (IS curve) with a positive EIS is the core of dynamic general equilibrium models used in the monetary policy analysis, e.g. New-Keynesian models. The sign and magnitude of the aggregate EIS determine the real effect of monetary policy. Therefore, the negative sign of the EIS changes key predictions of these models. Most notably, it implies that an increase in the real interest rate is expansionary.

I.A. Related literature

This paper is related to several areas in the finance and macro literature. First, numerous papers study the empirical properties of the real interest rate (Mishkin, 1981; Fama and Gibbons, 1982; Evans and Lewis, 1995; Ang, Bekaert, and Wei, 2007c). The overall message of these papers is that statistical properties of the real rate change over time. Similarly, the literature studying inflation documents that both the persistence and the volatility of inflation in the US vary over time (Barsky, 1987; Cogley, Primiceri, and Sargent, 2010). Despite this unambiguous evidence most of the recent studies constructing the ex-ante real interest rate assume linearity which leads to biased estimates. To avoid estimation issues, this paper proposes a measure of the ex-ante real rate that does not require estimating a statistical model for inflation.

Second, the empirical literature documents that the slope of the nominal yield curve predicts real output growth (Estrella and Hardouvelis, 1991; Plosser and Rouwenhorst, 1994; Ang, Piazzesi, and Wei, 2006) and consumption growth (Harvey, 1988). There is no consensus as to the economic mechanism behind this predictability. One way to explain the predictive power of the slope is monetary policy (Bernanke and Blinder, 1992), but other studies argue against this interpretation (Estrella and Hardouvelis, 1991; Plosser and Rouwenhorst, 1994). For a recent discussion see Benati and Goodhart (2008). In addition to the predictability of real output, Fama (1990) and Mishkin (1990) show that the slope predicts changes in inflation. I show that the slope tracks the variation in the ex-ante real rate and is unrelated to expected inflation, hence the relationship between the slope and aggregate consumption/output growth represents the intertemporal optimality condition that is consistent with a structural rather than monetary policy interpretation. Unlike most of the papers studying the predictive power of the slope, I study the sample 1875-2009 that includes a period in which the Fed was not operational. This helps distinguish monetary policy from structural interpretation.

Third, econometric studies estimating the EIS from aggregate data exploit the intertemporal optimality condition for consumption (Hansen and Singleton, 1982; Hall, 1988) or output (Fuhrer and Rudebusch, 2004; Bilbiie and Straub, 2012). Estimates of the EIS from aggregate data have sizeable standard errors and are usually close to zero. A variety of explanations for the low magnitude of the EIS estimates have been proposed: non-separability of non-durable and durable consumption (Ogaki and Reinhart, 1998), limited stock market participation (Guvenen, 2006), bounded rationality (Gabaix, 2012), or the structural break in the EIS (Bilbiie and Straub, 2012). These studies provide arguments for why the EIS estimates are biased toward zero but do not explain why the aggregate EIS is consistently negative. The lack of accurate EIS estimates from aggregates has motivated the analysis of household-level data. Vissing-Jorgensen (2002) and Attanasio, Banks, and Tanner (2002) show that the EIS is close to unity for stockholders. Similarly, Gruber (2006) estimates the EIS to be larger than one in a sample of households paying substantial capital income tax. My estimates of the EIS from the aggregate data are negative which makes the discrepancy between the micro and macro estimates of the EIS larger than previously documented. I argue that the differences arise as a result of aggregation of heterogenous households. The model presented in this paper is consistent with both the positive EIS for households participating in asset markets and the negative aggregate EIS.

Fourth, the evidence on violation of the permanent income hypothesis has led to models in which the consumption is generated by two types of consumers: forward-looking who consume their permanent income and rule-of-thumb who do not optimize intertemporally because they are excluded from asset markets (Campbell and Mankiw, 1989). Guvenen (2006) introduces heterogeneity in the EIS combined with limited stock market participation into a real business cycle model to replicate the empirical evidence on low (but positive) EIS at the aggregate level. I use the limited participation combined with nominal rigidities to study the implications of the negative aggregate EIS for dynamics of short-term interest rates.

Finally, asset pricing models usually represent the variation in the yield curve slope as a time-varying risk premium. Wachter (2006), Rudebusch and Swanson (2012), Bansal and Shaliastovich (2010) fall into this category. This paper provides an alternative interpretation of the yield curve slope. I first show that the slope closely tracks the variation in the real short-term interest rate. Then I use the estimated EIS to connect the variation in the short-term real rate to expected consumption growth.

II. The main result

This section states the main result and all the details are postponed to subsequent sections. I estimate the log-linearized version of the intertemporal optimality condition for aggregate consumption C_t using annual data in the sample 1875-2009:

$$\Delta c_{t+1} = \bar{\mu}_t + \sigma \tilde{r}_t + \epsilon_{t+1}, \tag{1.1}$$

where $\Delta c_{t+1} = \log C_{t+1} - \log C_t$, $\bar{\mu}_t$ collects constant terms and the time-varying second moments, σ denotes the EIS, $\tilde{r}_t$ is the measure of the ex-ante real interest rate proposed in this paper and ϵ_{t+1} subsumes expectations errors. I argue that, to obtain unbiased EIS estimates, the choice of instruments for the ex-post real rate r_{t+1} is of great importance. I exclude realized inflation from the instrument set because, as I will argue below, its forecast errors are predictable and the dynamics are non-linear. Instead, I use $\tilde{r}_t$ as an instrument for r_{t+1}.

Table I reports the estimates of σ across different sub-samples (estimation details are postponed to Section IV.A). The sub-samples are chosen so that they allow for an assessment of the stability of the estimates in different macroeconomic/policy environments and for comparison with the literature. Panel A reports the results for consumption and Panel B for the real output. The most important finding across sample periods and output measures is that the estimates are negative and statistically significantly different from zero.[3] The full

[3]Note that Hansen and Singleton (1984, 1996) in some cases obtain negative but insignificant estimates of the risk aversion parameter. Hall (1988) also estimates a negative value for σ using annual data in the sample spanning 1924-1940 and 1950-1983. Similarly, Yogo (2004) reports negative albeit in most cases insignificant estimates of EIS for a panel of developed countries. Using the US data on the long sample period (1891-

sample estimate of aggregate EIS is -0.51. The estimates provided in both panels of Table I counter the standard economic intuition which asserts that $\sigma > 0$, i.e. a positive shock to the real interest rate induces agents to postpone today's consumption to the next period.

In the subsequent sections, I elaborate on this result. Section III provides arguments for why $\tilde{r}_t$ is an appropriate measure of the ex-ante real interest rate. Section IV discusses the estimation details and econometric issues. Section II shows that the negative EIS estimates can be replicated in a standard general equilibrium model with limited asset market participation for plausible values of structural parameters. Section VI discusses the key mechanism for obtaining the negative aggregate EIS and studies its implications for term structure modeling. Section VII relates this paper's results to long-run risk models.

1997), Campbell (2003) estimates a negative though insignificant EIS parameter. These studies interpret negative estimates of EIS as implausible.

Table I: Estimates of Euler equation for total consumption and aggregate output, 1875-2009

The table reports the results for the log-linearized Euler equation given by (1.1). The IV setup is implemented as a two-step GMM. Panel A reports the results for consumption and Panel B for the real output. Each column corresponds to a different sample period. The first column shows the full sample estimates, 1875-2009. The second column reports the results for 1875-1979, a period preceding the Great Moderation and Volcker chairmanship. The third column displays the results for the post-war period 1947-2009. The fourth column reports the estimates for 1960-2009, which includes the inflationary period and finally the fifth column reports the results for 1980-2009, a post-inflation/Great Moderation period. The last column in both panels reports the estimates using the survey-based ex-ante real rate. Inflation expectations are obtained from the Livingston panel for the period 1960-2009. The data are annual. The autocorrelation and heteroskedasticity-robust standard errors are reported in parentheses (bandwith=3, Bartlett kernel). The overidentification test refers to Hansen's J statistic which is asymptotically χ^2_3 distributed. The "Weak identification test" in each panel reports the F-statistic (Kleibergen-Paap). The critical values for the maximum relative bias are 16.85 for 5%, 10.27 for 10%, 6.71 for 20%, and 5.34 for 30%. The critical values for the maximal size are 24.58 for 10%, 10.26 for 20%, and 8.31 for 25%. The critical values are computed using the procedure proposed by Stock and Yogo (2002). The last row in each panel reports the Wald test p-values testing the equality of σ in the pre-Volcker (1875-1979) and post-Volcker periods (1980-2009). The following variables are used as instruments: two lags of $\tilde{r}_t$, lagged three-month nominal T-bill and twice lagged consumption/output growth.

	1875-2009	1875-1979	1947-2009	1960-2009	1980-2009	1960-2009 r_t^{surv}
A. Consumption Euler equation						
$\widehat{\sigma}$	-0.51	-0.63	-0.72	-0.45	-0.69	-0.38
Robust s.e.	(0.20)	(0.25)	(0.28)	(0.27)	(0.24)	(0.28)
Confidence interval (95%)	[-0.91, -0.12]	[-1.12, -0.14]	[-1.27, -0.18]	[-0.97, 0.07]	[-1.17, -0.21]	[-0.93, 0.18]
Overidentification test (p-val)	0.78	0.91	0.17	0.16	0.21	0.24
Weak identification test (F-stat)	20.78	22.10	15.26	19.43	14.94	4.73
Wald test (p-val): σ equal	–	–	–	–	0.82	–
B. Output Euler equation						
$\widehat{\sigma}$	-0.51	-0.85	-1.14	-1.16	-1.08	-0.86
Robust s.e.	(0.29)	(0.41)	(0.33)	(0.33)	(0.28)	(0.34)
Confidence interval (95%)	[-1.07, 0.04]	[-1.65, -0.05]	[-1.79, -0.49]	[-1.81, -0.50]	[-1.63, -0.53]	[-1.53, -0.20]
Overidentification test (p-val)	0.60	0.69	0.23	0.60	0.73	0.42
Weak identification test (F-stat)	19.51	19.82	12.69	15.92	13.05	5.02
Wald test (p-val): σ equal	–	–	–	–	0.41	–

III. Measuring the real interest rate

The ex-ante short term real interest rate r_t is defined as $r_t = i_t - \mathrm{E}_t\pi_{t+1}$, where i_t is the short term nominal yield and $\mathrm{E}_t\pi_{t+1}$ represents the expected inflation one period ahead. While i_t is observable, r_t is not because the inflation expectations are not directly observable. There are two common ways to obtain inflation expectations.

The first and most often used method is to obtain r_t from the ex-post real rate $i_t - \pi_{t+1}$ by projecting it on the information set available at time t. Two key assumptions underlie this approach: (i) the rational expectations (RE) which puts a structure on the inflation forecast errors and (ii) the stability of the data-generating process for inflation, henceforth linearity assumption. However, empirical evidence does not square with either of these assumptions. As a consequence, measures of expected inflation based on realized inflation and linear regressions lead to biased estimates. More details on these assumptions are provided in Section III.B.

Second, the real rate r_t can be constructed using inflation survey data. The attractive feature of this approach is that a measure of r_t is obtained without the use of assumptions (i) and (ii). One concern related to the survey data is their accuracy. The empirical evidence documents, however, that survey-based inflation forecasts outperform statistical methods out-of-sample, often by a wide margin (Ang, Bekaert, and Wei, 2007a; Faust and Wright, 2012). Taken together, lack of accurate measures of the ex-ante real rate in long sample periods can explain imprecise estimates of the EIS from aggregate data.

This paper proposes an alternative measure of r_t denoted by $\tilde{r}_t$ that uses only the nominal yield curve data. Specifically, I define $\tilde{r}_t = i_t - y_t^{(n)}$ with n large. Appendix IX.B provides details on data used to obtain $\tilde{r}_t$. Constructing $\tilde{r}_t$ by subtracting the long-term nominal yield $y_t^{(n)}$ from i_t is based on the idea that most of the variation in long-term nominal yield is due to changing inflation expectations. In other words, I assume that the Fisher effect holds in the long-run. By this assumption, the real rate is less persistent than inflation expectations. This method has several advantages. First, $\tilde{r}_t$ relies neither on the RE or the linearity assumption, both of which are necessary when extracting inflation expectations from the realized inflation. That is to say, $\tilde{r}_t$ exploits the linear relationship between long-term yields and inflation expectations rather than modeling changing persistence and stochastic volatility of the inflation process. Second, given that the nominal Treasury yield curve data are available for much longer periods than inflation surveys, $\tilde{r}_t$ allows to study the real rate and its links to key macro variables over long periods. This is particularly important for the identification of the EIS.

To check its validity, Figure 1.1 compares $\tilde{r}_t$ to the ex-ante real rate obtained from the inflation survey denoted by r_t^{surv}. Inflation expectations are from the Livingston survey, the longest consistently available semi-annual inflation survey.[4]

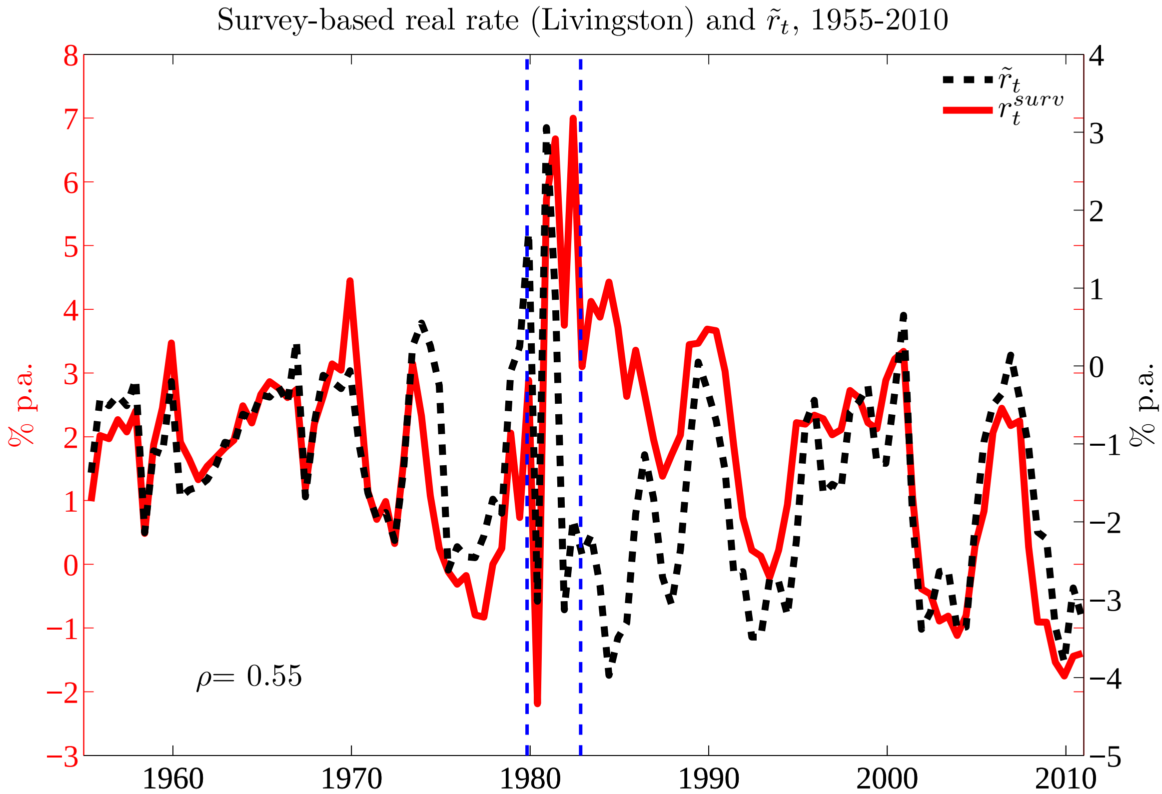


Figure 1.1: Survey-based ex-ante real interest rate and $\tilde{r}_t$, 1955-2010

The figure plots the measure of the ex-ante real rate $\tilde{r}_t$, which is computed as the difference between the three-month Treasury bill yield and ten-year Treasury yield together with the survey-based real interest rate $r_t^{surv} = i_t - \mathrm{E}_t^s \pi_{t+1}$, where i_t represents the three-month Treasury bill and $\mathrm{E}_t^s \pi_{t+1}$ denotes the survey-based measure of inflation expectations over the 12 months. The survey data are obtained from the Livingston survey with the median response being a proxy for the expectations. Dashed blue vertical lines mark the Volcker dis-inflation period October 1979 through October 1982 (M1 growth targeting period). The data are semiannual. The sample period is June 1955 through December 2010.

The most important observation is that $\tilde{r}_t$ closely tracks the variation in r_t^{surv} with the correlation $\rho = 0.55$. The only exception is a brief period at the beginning of 1980s, a period

[4] The Livingston survey is conducted in June and December every year. Participants are asked to forecast the CPI level six and twelve months ahead. Survey results are usually collected several weeks before the publication date. For a detailed discussion of issues related to the publication lag see Carlson (1977).

marked by extremely high and volatile inflation and interest rates due to the monetary policy experiment (Volcker disinflation).[5] Excluding the 1979-1984 period increases the correlation to 0.75. To formally show that $\tilde{r}_t$ is a valid instrument for the ex-ante real rate, I run the

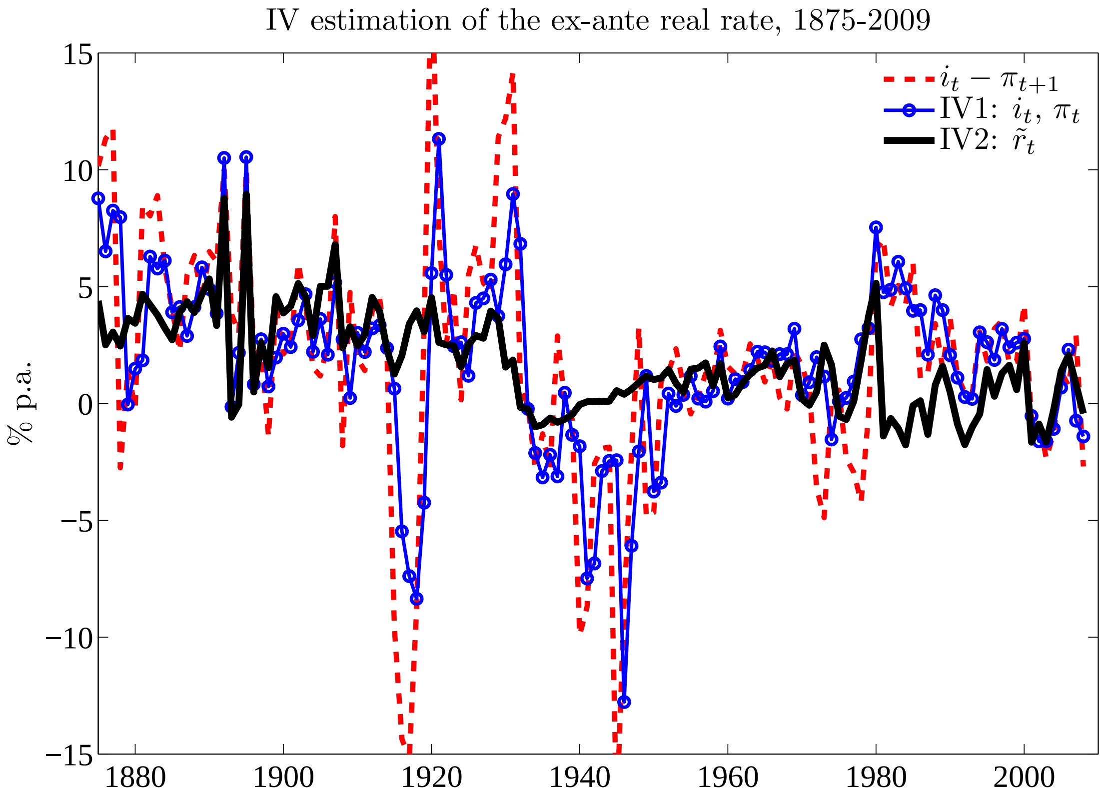


Figure 1.2: Instrumental variables estimation of the ex-ante real interest rate, 1875-2009

The figure compares two estimates of the ex-ante real rate, each obtained with a different set of instruments. First, the ex-post real rate $i_t - \pi_{t+1}$ (dashed red line) is regressed on the nominal short-term interest rate i_t represented by the three-month nominal T-bill and the past inflation π_t (blue line). Second instrument set includes only $\tilde{r}_t$ (thick black line). The dashed red line represents the ex-post real rate. The data are annual and un-smoothed. The sample period is 1875-2009.

following regression in the sample 1875-2009 (un-smoothed annual data):

$$i_t - \pi_{t+1} = \underbrace{\alpha_0}_{0.019\ [3.17]} + \underbrace{\alpha_1}_{1.056\ [4.79]} \tilde{r}_t + \varepsilon_{t+1} \quad \bar{R}^2 = 0.15, \tag{1.2}$$

[5] Designed to contain surging inflation, on October 6, 1979, the Fed formally announced a change in the conduct of monetary policy with the focus on reserves targeting. As a consequence, this change induced a large increase in the level and volatility of short-term nominal interest rates. In October 1982, the Fed abandoned the formal M1 growth target.

where the t-statistics reported in parentheses are Newey-West-adjusted with three lags. The regression results show that $\tilde{r}_t$ is indeed a valid instrument. Moreover, the null that $\alpha_1 = 1$ cannot be rejected, indicating a one-to-one mapping between $\tilde{r}_t$ and the ex-post real rate. Figure 1.2 plots the ex-ante real rate from the regression given by (1.2) together with the ex-post real rate. The close match of $\tilde{r}_t$ with the survey-based real rate (Figure 1.1) together with the regression results in (1.2) motivate the use of $\tilde{r}_t$ as a direct measure of the ex-ante real interest rate variation. Appendix IX.F provides additional results and discussion regarding the construction of $\tilde{r}_t$.

III.A. When is slope a good measure of the ex-ante real rate?

The idea of constructing a proxy for the real rate variation from the nominal yield curve builds on results in the earlier term structure literature (Fama, 1990; Mishkin, 1990). These papers document that the variation in real interest rates is concentrated at the short end of the yield curve and the slope is informative about the variation in the real rate. Cieslak and Povala (2011) show that the slope is highly correlated with the short-term nominal interest rate after extracting the long-term inflation expectations. Taken together, the empirical evidence locates most of the variation in slope at the short end of the real yield curve. Below, I discuss the conditions under which $\tilde{r}_t$ is an accurate measure of the variation in the short term real rate r_t. These assumptions impose weaker restrictions on the joint behavior of inflation and real rate than the implicit assumptions in parametric VAR models.[6]

Assumption 1: The ex-ante real interest rate is mean-reverting. Therefore the variation in long-term nominal yields due to the fluctuating real interest rate is negligible. The theoretical support for this assumption comes from the fact that in consumption-based models, the real rate and expected consumption growth rate should have similar time series properties. Arguably, the consumption growth rate is not exposed to permanent shocks. The empirical support for this assumption is the fact that long-term yields do not exhibit

[6]The idea is to combine the advantages of vector autoregressive models with time-varying parameters (TVP-VAR) and the parsimony of the constant parameter VAR. Therefore, the assumptions do not rule out shifts in parameters but also avoid the filtering of a large number of state variables as it is the case for TVP-VAR models.

large and persistent swings in periods characterized by a metallic monetary system such as Gold standard.[7]

Assumption 2: The variation in risk compensation accruing to long-term nominal bond holders is small in magnitude and has lower persistence than the real rate variation. Assuming RE, the empirical literature has identified time-varying risk premia as the main source of the rejection of the expectations hypothesis. At the same time, there is evidence from survey data showing significant departures from the full information rational expectations assumption in the short rate expectations (Froot, 1989; Piazzesi and Schneider, 2011; Cieslak and Povala, 2013). Therefore, deviations from RE are likely to explain part of the failure of the expectations hypothesis. Indeed, Cogley (2005) evaluates the joint RE–expectations hypothesis and finds that departures from the prescriptions of rational expectations are important for describing the observed yield curve dynamics. The time-varying risk premia extracted from the long term nominal Treasuries, despite high predictability of returns, have a half-life of less than a year and contribute around 5% to the total yield variation (Cieslak and Povala, 2011). In most of the analysis, I use annual data, therefore the variation in risk premia is unlikely to have a sizable impact on my results.

Assumption 3: Inflation expectations are described by a single factor with unit root persistence and stochastic volatility, i.e. I assume an unobserved component model with stochastic volatility (UC-SV) for realized inflation. This assumption is motivated by the observation that such a model provides a good statistical representation of inflation in the US (Stock and Watson, 2007) and in the UK (Cogley, Sargent, and Surico, 2012a). In the UC-SV model, inflation expectations are modeled as a single-factor random walk process with stochastic volatility. The assumption has several empirical implications. First, the term structure of inflation expectations is flat. Second, the real and nominal slope of the yield curve are identical. Third, short- and long-term nominal yields are cointegrated. The empirical evidence is largely consistent with all of these implications. Kozicki and Tinsley (2006) use inflation survey data from multiple sources to estimate the term structure of expected inflation. The estimated term structure is virtually flat throughout their sample period

[7] Rose (1988) cannot reject the null of unit root in the ex-post real rate. However, Garcia and Perron (1996) identify several regime switches in the post-war period and argue that the failure to reject the non-stationarity of the real interest rate is due to a small number of changes in mean rather than unit root. For measuring the real rate by $\tilde{r}_t$, it is important that the contribution of the real interest rate variation to the total variance of long term nominal yields is negligible. This continues to be the case even if the real rate undergoes a low number of regime switches.

starting in 1955. Additionally, the long-term survey data from the Blue Chip Economic Indicators (BCEI) panel for the 1984-2010 period, depicted in Panel a of Figure 1.3, confirm that the term structure of inflation expectations is indeed well described by a single level factor. Finally, Panel b of Figure 1.3 plots the real slope superimposed with the nominal slope for the period 1979-2010. Their correlation is 0.84.

For simplicity, the illustration below restricts the aforementioned assumptions to the case with fixed parameters and constant volatility. The one-period nominal interest rate can be written as:

$$i_t = r_t + \mathrm{E}_t \pi_{t+1}. \tag{1.3}$$

The n-period nominal Treasury yield can be decomposed as follows:

$$y_t^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} \mathrm{E}_t r_{t+i} + \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}_t \pi_{t+i} + rp_t^{(n)}, \tag{1.4}$$

where $rp_t^{(n)}$ represents the risk premium. Assume that the real rate follows an AR(1) process with zero mean: $r_t = \phi_r r_{t-1} + \sigma_r \varepsilon_t^r$, where $|\phi_r| < 1$, further assume that $\pi_t = \tau_t + \eta_t$ where inflation expectations τ_t are described by the random walk: $\tau_t = \tau_{t-1} + \sigma_\tau \varepsilon_t^\tau$ and η_t is a serially uncorrelated shock, then we have that $\tilde{r}_t = i_t - y_t^{(n)}$:

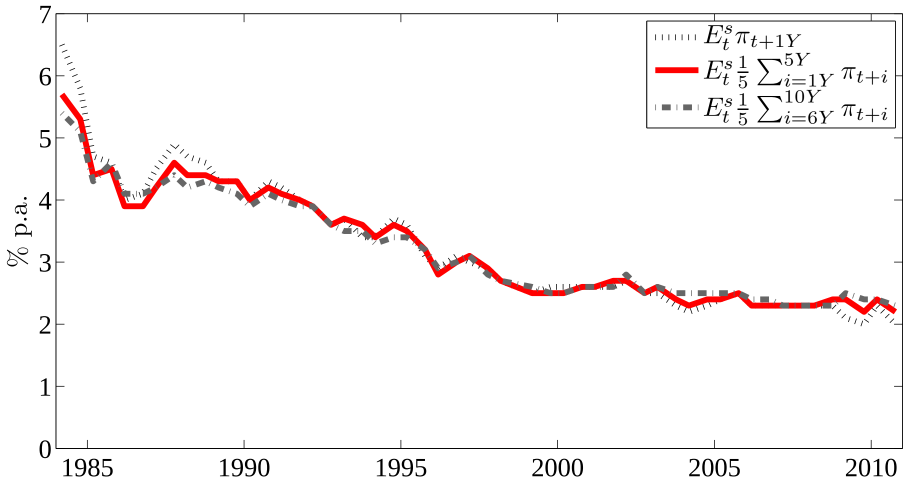
$$\begin{aligned} \tilde{r}_t &= r_t - \frac{1}{n} \sum_{i=0}^{n-1} \mathrm{E}_t r_{t+i} + \mathrm{E}_t \pi_{t+1} - \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}_t \pi_{t+i} - rp_t^{(n)} \\ &= \underbrace{\left(1 - \frac{1}{n}\left(\frac{1-\phi_r^n}{1-\phi_r}\right)\right)}_{\lim_{n\to\infty}=1} r_t - rp_t^{(n)}. \end{aligned}$$

From Assumption 1 it follows that, for large n, the variation in $\frac{1}{n}\sum_{i=0}^{n-1} \mathrm{E}_t r_{t+i}$ is negligible.[8] Clearly, $\tilde{r}_t$ is not informative about the unconditional mean of the real rate but this is inconsequential for the EIS estimation.[9] Assumption 2 implies that the variation in $rp_t^{(n)}$ is of smaller magnitude than the variation in the real rate, i.e. $\mathrm{Var}\left(r_t\right) \gg \mathrm{Var}\left(rp_t^{(n)}\right)$.

[8]Empirically, $\phi_r \approx 0.7$ at semi-annual frequency, which is obtained using the Livingston survey-based proxy for r_t in the 1955-2010 sample. For $y_t^{(n)}$, I use the ten-year yield, which implies $\left(1 - \frac{1}{n}\left(\frac{1-\phi_r^n}{1-\phi_r}\right)\right) \approx 0.83$.

[9]Indeed, one of the reasons for estimating the EIS from a linearized Euler equation is the fact that $\tilde{r}_t$ is not informative about the unconditional mean of the real rate.

a. Survey-based term structure of inflation expectations (BCEI), 1984-2010



b. Real and nominal yield curve slope, 1979-2010

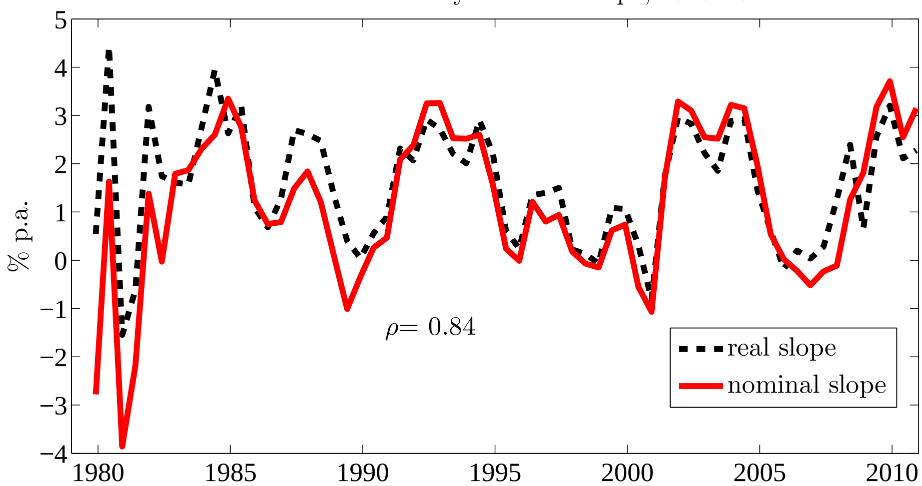


Figure 1.3: Term structure of inflation expectations

Panel *a* plots the survey-based inflation expectations obtained from the Blue Chip Economic Indicators survey. The data are semi-annual as the survey is conducted in March and October each year. The sample period is 1984 through 2010. $E^s_t \pi_{t+1Y}$ denotes the median survey response about the inflation one year ahead. $E^s_t \frac{1}{5} \sum_{i=1Y}^{5Y} \pi_{t+i}$ denotes the median survey response about the average inflation over the next five years and finally $E^s_t \frac{1}{5} \sum_{i=6Y}^{10Y} \pi_{t+i}$ represents the median response about the average inflation between five and ten years ahead. Panel *b* compares the real and nominal slope of the yield curve. The data are semi-annual. The sample period is 1979-2010. The start of the sample is dictated by the availability of the long term inflation survey data. The real slope is constructed as the difference of the ten-year real yield and real three-month Treasury bill. The real ten-year yield is obtained using the ten-year inflation expectations from Livingston and SPF surveys.

$\mathrm{E}_t\pi_{t+1} - \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}_t\pi_{t+i} = 0$ which follows from Assumption 3. If Assumptions 1-3 hold, $\tilde{r}_t$ or equivalently the yield curve slope provides an accurate measure of the real rate variation.

III.B. The RE and linearity assumptions

It is instructive to study the two key assumptions that underlie the standard method for obtaining the ex-ante real rate.

RE assumption. The RE assumption asserts that inflation forecast errors are not predictable. If the forecast errors are predictable, expected inflation obtained from standard time-series models will inherit the predictable part of these errors. The slope of the term structure of nominal yields varies because of changes in expected inflation and fluctuations in the real rate (Fama, 1990). Additionally, inflation and the real interest rate are negatively correlated (Mishkin, 1981; Fama and Gibbons, 1982). These stylized facts are based on the following regression:

$$\pi_{t+m} - \pi_{t+1} = \delta_0 + \delta_1 \mathrm{slope}_t + \varepsilon_{t+m}, \tag{1.5}$$

where π_{t+m} denotes the realized inflation between time t and $t+m$, $m > 1$ and π_{t+1} represents the realized inflation one period ahead.

Panel A of Table II reports the results for the regression (1.5) for $m = 2, \ldots, 5$ years in the sample 1955-2010. The results confirm that slope is a significant predictor of realized inflation changes for all horizons considered with R^2's up to 20% and the positive sign. Estimates of δ_1 are significantly below unity which indicates that the slope also predicts the real rate variation. To understand the source of the inflation predictability, I decompose the dependent variable in (1.5) as follows:

$$\pi_{t+m} - \pi_{t+1} = \pi_{t+m} - \mathrm{E}_t\pi_{t+1} - \underbrace{(\pi_{t+1} - \mathrm{E}_t\pi_{t+1})}_{\text{inflation forecast error}},$$

where the Livingston inflation survey data are used to proxy for $\mathrm{E}_t\pi_{t+1}$. Panels B and C of Table II report the regression results for $(\pi_{t+m} - \mathrm{E}_t\pi_{t+1})$ and $(\pi_{t+1} - \mathrm{E}_t\pi_{t+1})$ as dependent variables, respectively. The regressions indicate that all of the predictability of realized inflation is due to the predictability of one-period inflation forecast errors. This result has

Table II: Inflation changes and the slope, 1955-2010

Panel A reports the results for the regression (1.5) for different horizons $m = \{2, \ldots, 5\}$ years. Using inflation surveys, the dependent variable in (1.5) is decomposed as follows: $\pi_{t+m} - \pi_{t+1} = \pi_{t+m} - \mathrm{E}_t^s \pi_{t+1} - (\pi_{t+1} - \mathrm{E}_t^s \pi_{t+1})$ where $\pi_{t+1} - \mathrm{E}_t^s \pi_{t+1}$ is the one-year inflation forecast error. $\pi_{t+m} = (\log P_{t+m} - \log P_t)/m$ where P_t is the CPI index. Panels B and C report the regression results for the decomposed dependent variable. $\mathrm{E}_t^s \pi_{t+1}$ is obtained from the Livingston survey which is conducted semi-annually (June and December). Timing of the variables is as follows. For example, a June 1955 slope observation is constructed using end of May 1955 bond data. A June 1955 observation for the one-year inflation rate is computed from the May 1955 and May 1956 CPI. The sample covers the period 1955-2010. Both the sampling frequency and beginning of the sample period are determined by the availability of inflation survey data. Newey-West adjusted t-statistics are reported in parentheses, the number of lags is 6.

	m=2	m=3	m=4	m=5
Panel A. $\pi_{t+m} - \pi_{t+1} = \delta_0 + \delta_1 \mathbf{slope}_t + \varepsilon_{t+m}$				
δ_1	0.28	0.48	0.58	0.56
	(2.97)	(3.04)	(2.80)	(2.43)
$\bar{R}^2$	0.15	0.20	0.20	0.16
Panel B. $\pi_{t+m} - \mathbf{E}_t^s \pi_{t+1} = \beta_0 + \beta_1 \mathbf{slope}_t + \varepsilon_{t+1}$				
δ_1	-0.22	-0.04	0.04	0.03
	(-1.32)	(-0.20)	(0.16)	(0.11)
$\bar{R}^2$	0.02	0.00	0.00	0.00
Panel C. $\pi_{t+1} - \mathbf{E}_t^s \pi_{t+1} = \gamma_0 + \gamma_1 \mathbf{slope}_t + \varepsilon_{t+1}$				
γ_1	-0.49			
	(-3.27)			
$\bar{R}^2$	0.14			

several important implications. First, slope is not informative about the variation in expected inflation which supports the notion that slope predominantly captures the variation in the real interest rate.[10] Second, it indicates a violation of the RE assumption for inflation.[11] As a consequence, the real rate obtained from the ex-post realized real rate contains the predictable part of inflation forecast errors which leads to more volatile and biased estimates of the ex-ante real rate. Third, the negative correlation between the real rate and the expected inflation reported in the literature disappears once the ex-ante real rate is obtained

[10] Equivalently, yield curve slope does not predict changes in survey-based inflation expectations (Livingston).

[11] I interpret the deviations from the RE assumption broadly. For instance, these can include a peso problem, i.e. it can well be that investors anticipated a discrete shift in the inflation process which was then reflected in asset prices long before it actually occurred. Alternatively, the predictable forecast errors can arise in a setup with imperfect information. From the perspective of monetary policy, predictable inflation forecast errors are indicative of the inflation non-neutrality in the short-run. To explicitly account for the predictable inflation forecast errors documented above, it is straightforward to extend the UC-SV model by specifying an AR(1) dynamics for the transitory component.

from the inflation survey data. In the 1955-2010 sample, the correlation between the ex-ante real rate and inflation expectations is 0.16.

Linearity assumption. A number of papers document that the statistical properties of inflation change over time (Barsky, 1987; Stock and Watson, 2007; Cogley, Sargent, and Surico, 2012b). Barsky (1987) shows that the inflation evolved from a highly volatile white noise process in the period before World War I into a highly persistent one in the post-1960 period. Stock and Watson (2007) show that the postwar US inflation is well described by an UC-SV model.[12] To evaluate the bias introduced by assuming the linear dynamics for inflation in the presence of stochastic volatility, I simulate inflation from the UC-SV model and the real rate as a simple AR(1) process using parameter values consistent with the data. Details of the simulation exercise are provided in Appendix IX.A. The median correlation between the true ex-ante real rate and the rate obtained by a linear projection is 0.34. In the simulation exercise, biased estimates of the real rate translate into a sizable bias in the EIS estimates. For the true value of EIS=1, the median estimate of the EIS from the linearized Euler equation is 0.30. The non-linearity can be handled by modeling the nominal short rate i_t and inflation π_t within a TVP-VAR with stochastic covariance matrix. Indeed, Cogley and Sargent (2005) show that models with time-varying parameters perform well under both random walk time variation and discrete structural breaks. However, a typical TVP-VAR with two lags and three observables (inflation, unemployment, nominal short rate) requires the filtering of more than 20 time-varying parameters, which necessarily introduces a significant degree of statistical uncertainty.

The degree to which the RE and linearity assumptions are violated can be seen by comparing the ex-ante real rate instrumented by $\tilde{r}_t$ and the ex-ante real rate obtained by regressing $i_t - \pi_{t+1}$ on i_t and π_t. Figure 1.2 plots the ex-post real rate together with these two versions of ex-ante real rate. It is clear that including past inflation improves the statistical fit, in fact the $\bar{R}^2$ increases from 0.15 obtained in regression (1.2) to 0.51. However, the plot shows that most of the improvement in statistical fit comes from fitting extreme values of inflation during the two World Wars and the Great Depression. Despite the fact that inflation is clearly a valid instrument in a statistical sense, it is not a valid instrument economically because the RE and linearity assumptions are violated. Figure 1.2 also shows that in the

[12] Earlier work by Evans and Lewis (1995) argues that a two-component model of inflation where the volatility of shocks to each component follows a regime-switching Markov process closely replicates the inflation expectations from the Livingston survey.

post-1980 period, the ex-post real rate and both measures of ex-ante real rate increasingly move together. Indeed, post-1980, the correlation between the slope and inflation forecast errors has been declining.[13]

IV. Estimating the EIS from the aggregate data

Assuming an isoelastic period utility function and the absence of market frictions, the intertemporal optimality condition for consumption C_t reads:

$$\mathrm{E}_t \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + R_{rf,t+1}) \right] = 1, \tag{1.6}$$

where γ denotes the risk aversion parameter, β represents the time discounting and $R_{rf,t+1}$ is the ex-post real interest rate. Further, assuming conditional log-normality leads to a linearized version of (1.6):

$$\Delta c_{t+1} = \mu_t + \frac{1}{\gamma} r_t + u_{t+1}, \tag{1.7}$$

where $r_t = \mathrm{E}_t \log(1 + R_{rf,t+1})$, u_{t+1} denotes the expectational error and:

$$\mu_t = \frac{\log(\beta)}{\gamma} + \frac{\gamma}{2} \mathrm{var}_t (\Delta c_{t+1}) - \mathrm{cov}_t (\Delta c_{t+1}, \log(1 + R_{rf,t+1})) + \frac{1}{2\gamma} \mathrm{Var}_t (\log(1 + R_{rf,t+1})).$$

In the previous section, I argue that r_t can be replaced by $\tilde{r}_t$ after accounting for the unconditional mean of the real rate. Therefore, an estimable version of (1.7) reads:

$$\Delta c_{t+1} = \bar{\mu}_t + \sigma \tilde{r}_t + u_{t+1}, \tag{1.8}$$

where $\sigma \equiv \frac{1}{\gamma}$ and $\bar{\mu}_t$ is μ_t plus the unconditional mean of r_t.[14]

---

[13] Note that all estimates of the EIS using household-level data (e.g. CEX survey) fall into the post-1980 period where the linear projection method for obtaining the ex-ante real rate works reasonably well, which explains the significant and plausible estimates of the EIS for households participating in asset markets.

[14] Note that due to the log-linearization, the isoelastic specification of the utility function is not crucial for relating the expected consumption growth to the real interest rate. Alternative utility specifications such as Epstein-Zin utility obtain the same relationship.

IV.A. Estimation

I use the US real per capita annual data on consumption and output for the sample period 1875-2009. I estimate the consumption and output Euler equation on two data sets. First, I use the data set constructed by Barro and Ursua (1875-2009). Second, the consumption data from the Bureau of Economic Analysis (BEA) are used at annual frequency (1929-2009) and at quarterly frequency (1950-2009). More details regarding both data sets and their sources are provided in Appendix IX.B. The output data allow me to evaluate the aggregate output Euler equation (IS equation) used in the New-Keynesian model below. The aggregate output Euler equation is the core of dynamic equilibrium models used in monetary policy analysis. Therefore, the EIS estimates from the output Euler equation are key for evaluating the real effects of monetary policy. Figure 1.4 plots the real consumption Δc_{t+1} (Panel a) and output growth Δy_{t+1} (Panel b) superimposed with $\tilde{r}_t$ for the period 1875-2009. The unconditional correlation between $\Delta c_{t+1}, \Delta y_{t+1}$ and $\tilde{r}_t$ is -0.34 and -0.30, respectively.

If one is willing to assume homoscedasticity and a household's decision interval of one year (in addition to Assumptions 1-3 from Section III.A), the log-linearized Euler equation given in (1.8) can be estimated via OLS. To account for possible correlation between $\tilde{r}_t$ and conditional variance terms collected in $\bar{\mu}_t$ and also to avoid potential issues induced by time-aggregation, I estimate (1.8) via instrumental variables. The linear IV estimation is standard and is implemented as a two-step linear GMM. The set of instruments for $\tilde{r}_t$ includes twice lagged consumption growth, two lags of $\tilde{r}_t$ and the lagged nominal three-month Treasury bill yield. The lags of instruments for $\tilde{r}_t$ are chosen such that there is no overlap with the consumption at time t.

IV.B. Estimation results

In addition to the main results reported in Table I based on Barro and Ursua's data, this section discusses the estimates of EIS from the BEA consumption data. I evaluate if the estimated magnitude of the EIS from annual data on total consumption matches the estimates from non-durable consumption and quarterly data, respectively.

Barro and Ursua's data set does not distinguish between non-durable and durable consumption expenditures due to limitations on data availability. Given that expenditures on durable goods are roughly 15% of the non-durable consumption and services, the potential bias in the

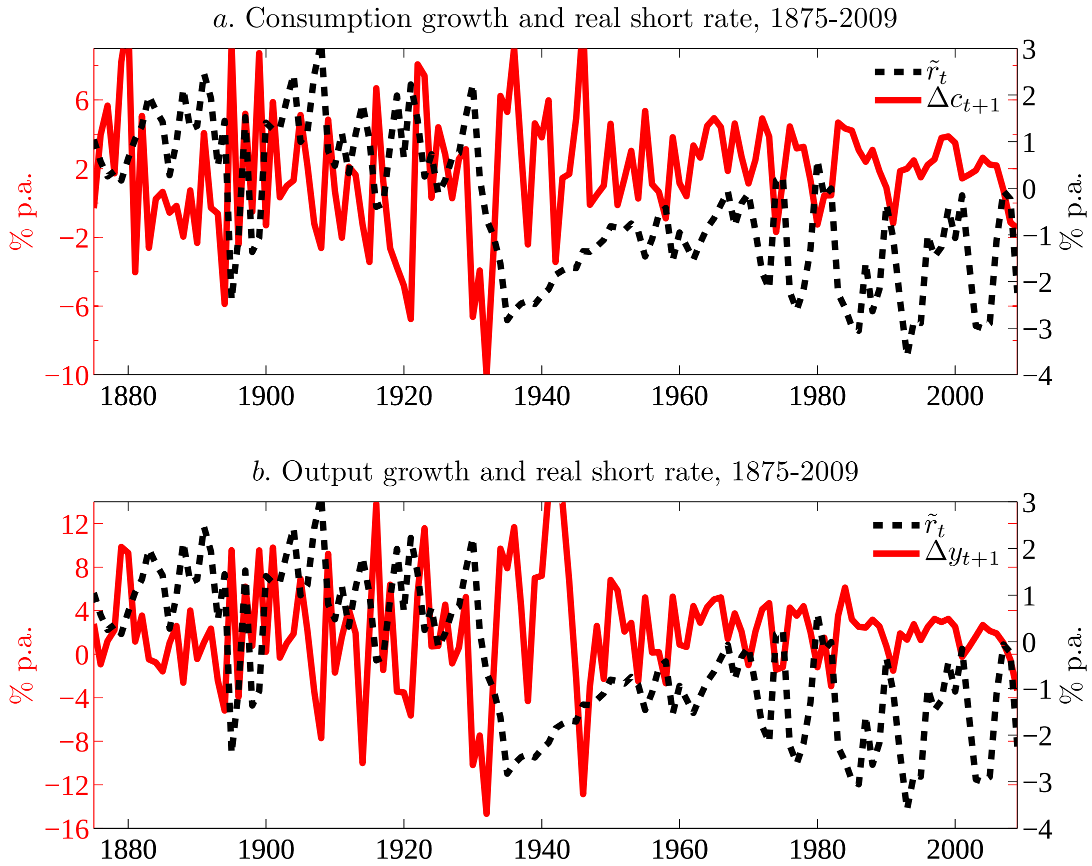


Figure 1.4: Consumption and output growth and the real interest rate, 1875-2009

Panel a plots real per capita consumption growth superimposed with the proxy for the short term real interest rate denoted by $\tilde{r}_t$. Panel b plots the real output growth superimposed with $\tilde{r}_t$. The data are annual and the sample period is 1875 through 2009.

EIS estimates is likely to be small.[15] To quantify the impact of durable goods expenditures on the EIS estimates, I estimate the consumption Euler equation using the non-durable consumption and services in the sub-sample 1929-2009. The sample start is determined by the availability of the data on non-durable consumption and services from the BEA. Table III reports the results for three different sample periods. Panel A displays the estimates of EIS using expenditures on non-durable goods and services. Panel B shows the results obtained

[15]The main motivation for excluding expenditures on durables is to minimize the measurement error originating from the discrepancy between expenditures and the consumption itself. However, with decreasing measurement frequency the error becomes smaller.

with durable goods expenditures and finally Panel C reports the EIS estimates using total consumption.

Table III: Euler equation estimates for non-durable consumption, 1929-2009

The table reports the estimation results for the log-linearized Euler equation given by (1.8). The IV setup is implemented as a two-step GMM. Panel A reports the results for non-durable consumption and services, Panel B for the durable consumption expenditures and Panel C for the total consumption. Each column corresponds to a different sample period. The first column shows the full sample estimates, 1929-2009. This is the longest sample period for which the split into durable and non-durable consumption expenditures is available. The second column reports the results for 1947-2009, a post-war period. The third column displays the results for the period 1980-2009. The data are annual. The autocorrelation- and heteroskedasticity-robust standard errors are reported in parentheses (bandwith=3, Bartlett kernel). The overidentification test refers to Hansen's J statistic which is asymptotically χ^2_3 distributed. The "Weak identification test" in each panel reports the F-statistic (Kleibergen-Paap). The critical values for the maximum relative bias are 16.85 for 5%, 10.27 for 10%, 6.71 for 20% and 5.34 for 30%. The critical values for the maximal size are 24.58 for 10%, 10.26 for 20%, and 8.31 for 25%. The critical values are computed using the procedure proposed by Stock and Yogo (2002). The following variables are used as instruments: two lags of $\tilde{r}_t$, the lagged three-month nominal T-bill and twice lagged corresponding consumption growth.

	1929-2009	1947-2009	1980-2009
A. Nondurable consumption & services			
$\widehat{\sigma}$	-0.66	-0.47	-0.45
Robust s.e.	(0.30)	(0.21)	(0.17)
Confidence interval (95%)	[-1.25, -0.07]	[-0.88, -0.05]	[-0.79, -0.12]
Overidentification test (p-val)	0.41	0.69	0.25
Weak identification test (F-stat)	21.77	14.47	15.67
B. Durable expenditures			
$\widehat{\sigma}$	-2.43	-2.59	-2.52
Robust s.e.	(0.88)	(0.74)	(0.79)
Confidence interval (95%)	[-4.14, -0.71]	[-4.05, -1.13]	[-4.07, -0.96]
Overidentification test (p-val)	0.90	0.29	0.19
Weak identification test (F-stat)	18.30	11.73	12.00
C. Total consumption			
$\widehat{\sigma}$	-0.78	-0.68	-0.64
Robust s.e.	(0.32)	(0.24)	(0.21)
Confidence interval (95%)	[-1.40, -0.15]	[-1.14, -0.22]	[-1.05, -0.23]
Overidentification test (p-val)	0.43	0.58	0.20
Weak identification test (F-stat)	22.03	14.46	16.48

The estimate of EIS using non-durables in the 1929-2009 period is -0.66 and statistically significant. It compares to the estimate of -0.78 using total consumption in the same period. Similarly, the EIS estimates using non-durable and total consumption expenditures in the 1980-2009 period are -0.45 and -0.64, respectively. Panel B shows that the interest rate elasticity of durable expenditures is a multiple of the non-durable consumption sensitivity. Higher sensitivity of durable goods expenditures slightly increases the magnitude of EIS estimates when total consumption is used. However, the difference is well within one standard deviation of the EIS estimate from the non-durable consumption. The estimates are stable across sub-samples. Particularly, in the most recent period 1980-2009, the estimates of aggregate EIS have relatively low standard errors despite the low number of observations. Overall, there is no evidence that the negative aggregate EIS is driven solely by the non-durable consumption expenditures.

The estimates of σ in Tables I and III are unaffected by the structural break in inflation and monetary policy which is usually located in the 1979-1982 period (Clarida, Gali, and Gertler, 2000; Bilbiie and Straub, 2012).[16] Comparing the second and fifth columns in both panels of Table I shows neither a change of sign nor a significant difference in the magnitude of $\hat{\sigma}$. A formal test of equality of $\hat{\sigma}$ in the period 1875-1979 and 1980-2009 is reported in the last row of each panel. In both cases the null that the coefficients are equal is not rejected.[17]

Throughout the paper, I use annual consumption and interest rate data. An important question is if the results extend to quarterly frequency. One concern is that if the household's decision interval is shorter than one year, the aggregation introduces a spurious autocorrelation in Δc_{t+1} and the estimates of the EIS are biased. Main arguments for measuring the consumption and output at annual rather than higher frequencies are the transaction costs and other frictions which might prevent households from instantaneous consumption smoothing (Gabaix and Laibson, 2001; Jagannathan and Wang, 2007). Additionally, the variation in real interest rates is persistent (half-life above one year), hence the consumption could react slowly to changes in the real interest rate because the utility loss induced by the slow adjustment is low. To evaluate whether my results are robust to time aggregation, I re-estimate the consumption Euler equation on quarterly data for the period 1950-2009. To obtain estimates of the EIS that are comparable to the annual data, quarterly consumption growth is annualized. Table IV reports the estimates of σ for non-durables (Panel A), durable consumption expenditures (Panel B) and the total consumption (Panel C) across

[16]Bilbiie and Straub (2012) argue that the main reason for the structural break and the change of sign in σ is the financial market liberalization which stimulated an increase in asset market participation. However, they use realized inflation when constructing the ex-ante real rate which is responsible for the identified structural break.

[17]In fact, in other countries that have not experienced a structural break in inflation in the recent past (e.g. Germany, the Netherlands or Canada), the standard method yields negative and significant estimates of the aggregate EIS of the similar magnitude compared to the US estimates using $\tilde{r}_t$.

Table IV: Euler equation estimates for quarterly data, 1950-2009

The table reports the estimation results for the log-linearized Euler equation given by (1.8). The IV setup is implemented as a two-step GMM. Panel A reports the results for non-durable consumption and services, Panel B for the durable consumption expenditures and Panel C for the total consumption. Quarterly consumption growth data are annualized to be comparable with $\tilde{r}_t$ which is annual. Each column corresponds to a different sample period. The first column shows the full sample estimates for which the quarterly consumption data are available, 1950-2009. The second column reports the results for 1960-2009. The third column displays the results for the period 1980-2009. The data are quarterly. The autocorrelation- and heteroskedasticity-robust standard errors are reported in parentheses (bandwith=8, Bartlett kernel). The overidentification test refers to Hansen's J statistic which is asymptotically χ^2_5 distributed. The "Weak identification test" in each panel reports the F-statistic (Kleibergen-Paap). The critical values for the maximum relative bias are 19.28 for 5%, 11.12 for 10%, 6.76 for 20%, and 5.15 for 30%. The critical values for the maximal size are 29.18 for 10%, 11.72 for 20%, and 9.38 for 25%. The critical values are computed using the procedure proposed by Stock and Yogo (2002). Following variables are used as instruments: lag one and lag four of $\tilde{r}_t$, lags two through four of consumption growth, and lag one of the three-month nominal T-bill.

	1950-2009	1960-2009	1980-2009
A. Nondurable consumption & services			
$\widehat{\sigma}$	-0.24	-0.40	-0.50
Robust s.e.	(0.18)	(0.19)	(0.18)
Confidence interval (95%)	[-0.60, 0.12]	[-0.77,-0.03]	[-0.86, -0.15]
Overidentification test (p-val)	0.08	0.12	0.26
Weak identification test (F-stat)	217.55	254.22	151.09
B. Durable expenditures			
$\widehat{\sigma}$	-2.80	-2.64	-2.67
Robust s.e.	(0.81)	(0.80)	(0.77)
Confidence interval (95%)	[-4.33, -1.28]	[-4.13, -1.14]	[-4.18, -1.15]
Overidentification test (p-val)	0.35	0.51	0.60
Weak identification test (F-stat)	200.82	173.01	89.95
C. Total consumption			
$\widehat{\sigma}$	-0.35	-0.50	-0.62
Robust s.e.	(0.20)	(0.22)	(0.24)
Confidence interval (95%)	[-0.74, 0.04]	[-0.93, -0.07]	[-1.08, -0.15]
Overidentification test (p-val)	0.04	0.09	0.17
Weak identification test (F-stat)	206.22	220.89	124.34

three post-war periods. Throughout the table, the magnitudes of EIS closely correspond to the estimates from annual data (Table III) which indicates that the estimated magnitudes are invariant to the decision interval of households.

It is possible that the negative EIS is specific to the US data. To evaluate this interpretation, Table V reports the EIS estimates for selected developed countries. The EIS estimates are consistently negative and have, in some cases, a larger magnitude than in the US. Arguably, each of these countries follow different monetary policies. Therefore, the evidence from a

Table V: Euler equation estimates for selected developed countries

The table reports estimation results for the log-linearized Euler equation given by (1.8). The IV setup is implemented as a two-step GMM. Panel A reports the results for consumption and Panel B for the real output. Each column corresponds to a different country. The sample period for each country is 1950-2009, the only exception is Germany with the sample period 1956-2009 (due to the data availability). The data are annual and their sources are described in Appendix IX.B. The autocorrelation and heteroskedasticity-robust standard errors are reported in parentheses (bandwith=3, Bartlett kernel). The overidentification test refers to Hansen's J statistic which is asymptotically χ^2_3 distributed. The "Weak identification test" in each panel reports the F-statistic (Kleibergen-Paap). The critical values for the maximum relative bias are 16.85 for 5%, 10.27 for 10%, 6.71 for 20%, and 5.34 for 30%. The critical values for the maximal size are 24.58 for 10%, 10.26 for 20%, and 8.31 for 25%. The critical values are computed using the procedure proposed by Stock and Yogo (2002). The instruments are twice lagged consumption/output growth, two lags of $\tilde{r}_t$, and the lagged nominal Treasury bill yield.

	Canada	Germany	Netherlands	Switzerland	United Kingdom
A. Consumption Euler equation					
$\widehat{\sigma}$	-0.93	-0.96	-0.91	-0.28	-0.19
Robust s.e.	(0.21)	(0.34)	(0.28)	(0.22)	(0.24)
Confidence interval (95%)	[-1.35, -0.51]	[-1.62, -0.30]	[-1.46, -0.37]	[-0.70, 0.15]	[-0.66, 0.27]
Overidentification test (p-val)	0.30	0.08	0.39	0.33	0.70
Weak identification test (F-stat)	11.00	22.79	38.06	12.69	27.05
B. Output Euler equation					
$\widehat{\sigma}$	-0.95	-1.02	-0.84	-1.15	-0.41
Robust s.e.	(0.24)	(0.29)	(0.23)	(0.29)	(0.18)
Confidence interval (95%)	[-1.42, -0.49]	[-1.58, -0.46]	[-1.30, -0.38]	[-1.72, -0.59]	[-0.76, -0.06]
Overidentification test (p-val)	0.55	0.66	0.34	0.53	0.30
Weak identification test (F-stat)	10.00	22.84	33.44	11.92	30.79

panel of countries supports the structural interpretation of the negative EIS. Discussion of some additional econometric issues is provided in Appendix IX.C.

IV.C. Econometric issues

Estimating the EIS parameter from the log-linearized Euler is subject to a number of potential biases. One obvious source is the stochastic volatility which has been mostly studied at the household level (Carroll, 2001; Attanasio and Low, 2004). Attanasio and Low (2004) show that, even in the presence of stochastic volatility, the log-linearized Euler equation yields unbiased estimates of the EIS if the sample period is long enough, which is the case in this paper. Furthermore, the stochastic volatility is less of a concern in the aggregate data where the idiosyncratic volatility is averaged out. Beeler and Campbell (2008) show that the bias due to the stochastic volatility does not play an important role, at least for the volatility process calibrated to the US aggregate consumption data. Additionally, in Section

VII below, I show that the proxy for stochastic volatility of consumption growth does not significantly alter the EIS estimates.

The observation that the linearized Euler equation (1.7) and its reversed form with the real rate as a dependent variable lead to very different estimates of the EIS is commonly interpreted as an evidence for weak instruments (Neely, Roy, and Whiteman, 2001; Yogo, 2004). This is not the case for $\tilde{r}_t$. Using the identical instrument set, the estimate of $\frac{1}{\sigma}$ from the reversed form of the consumption Euler equation (1.8) is -1.66 with the 95% confidence interval: $[-2.65, -0.66]$ in the 1875-2009 sample. The full sample estimate of σ from the original specification is -0.51 which is within the confidence interval given by the reversed form.

I construct moving averages of both quantities, demeaned Δc_{t+1} and $\tilde{r}_t$, and run a univariate regression with filtered series. The goal of this exercise is to evaluate the relationship between consumption growth and the real rate across frequencies. The real rate variation is persistent and it should be reflected in the persistent variation of expected consumption growth if the EIS is non-zero. Conversely, if the significant estimates of the EIS reported above are driven by a few outliers, the filtered series shall not have a similar degree of co-movement. To compute moving averages, I use the filter proposed in Lucas (1980). For a covariance stationary series x_t, it is given by:

$$\bar{x}_t(\beta) = \alpha \sum_{k=-n}^{n} \beta^{|k|} x_{t+k} \tag{1.9}$$

$$\alpha = \frac{(1-\beta)^2}{1-\beta^2 - 2\beta^{n+1}(1-\beta)}, \tag{1.10}$$

where $\beta \in [\,0, 1\,)$. I study the following four degrees of smoothing $\beta \in \{0, 0.8, 0.95, 0.98\}$.[18]

Panels $a-d$ of Figure 1.5 evaluate the comovement of consumption growth with the real rate for different frequencies as measured by values of β. As the figure shows, the negative relationship between $\tilde{r}_t$ and consumption growth is present at all frequencies. Moreover, the regression coefficient, which is given by the slope of the dash-dotted line, shows a remarkable stability for all degrees of smoothing. The estimated slope coefficients are close to the EIS

[18]Note that $\beta = 0$ corresponds to a univariate regression of unfiltered series. For $\beta \to 1$, the regression approximately measures the relationship at frequency $\omega = 0$ in the frequency domain, for details see Whiteman (1984).

estimates reported in Panel A of Table I. The close relationship between the consumption growth and the real interest rate indicates that expected consumption growth contains a time-varying and persistent component. Section VII studies this relationship in greater detail and links the results to the long-run risk literature.

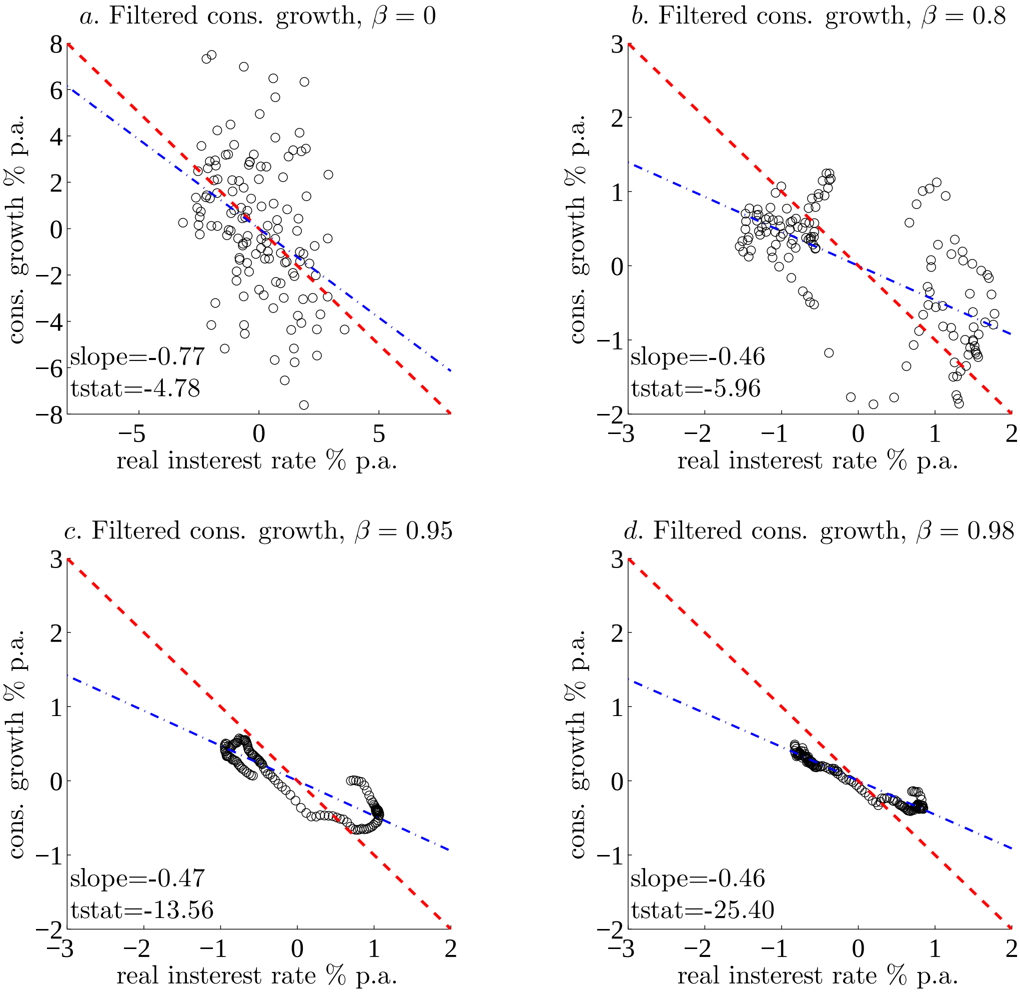


Figure 1.5: Consumption growth and real interest rate–Lucas filter, 1875-2009

Panels *a*–*d* scatter-plot the filtered consumption growth and the real rate measure $\tilde{r}_t$ for different filtering parameters β. For $\beta = 0$ (Panel *a*), the filter corresponds to a linear regression of unfiltered series. The filter is given by equation (1.9). I set $n = 50$. Blue dash-dotted lines represent the regression coefficient obtained from a regression of the filtered consumption growth on filtered $\tilde{r}_t$. The regression coefficients together with Newey-West adjusted t-statistics are reported in the lower left corner of each panel. The red dashed line represents the regression slope of minus one. The data are annual and the sample period is 1875 through 2009.

IV.D. Alternative explanations

The relationship between the real interest rate and consumption growth can potentially be negative for other reasons than the negative aggregate EIS. For instance, the time discounting can be time-varying and strongly negatively correlated with the real rate.[19] In such a case, one obtains a negative estimate of σ even though the true EIS is positive. However, for this to happen, the discount factor shock would have to be more volatile than r_t, which, together with the highly negative correlation seems implausible. Alternatively, it can be the case that the unobserved returns to human capital are negatively correlated with the real short rate. However, Lustig, Van Nieuwerburgh, and Verdelhan (2012) show that returns on human wealth are closely positively related to real bond returns. Therefore, the negative relationship between consumption growth and the real interest rate cannot be replicated for positive values of σ at the aggregate level unless one is willing to accept counterfactual assumptions.

V. Model

This section outlines a standard New Keynesian model with limited asset market participation. The model setup is closely related to Gali, Lopez-Salido, and Valles (2004) and Bilbiie (2008). While these papers focus on the equilibria determinacy in the presence of limited market participation, I evaluate the asset pricing implications of limited participation.[20] The key ingredient of the model is limited asset market participation rather than a short-run inflation non-neutrality induced by sticky prices. To make the key mechanism transparent, I restrict the technology to be the only economic disturbance in the model. The main purpose of the model is to illustrate one potential mechanism that can replicate the negative sign and the magnitude of the aggregate EIS. As such, the model is too simple to be evaluated on quantitative predictions along other dimensions, but could be easily enriched.

[19] The discount factor shock is often introduced in the empirical implementations of dynamic stochastic general equilibrium models (Justiniano and Primiceri, 2008). Given the empirical failure of the consumption Euler equation, the time-varying discounting picks up the variation in consumption growth. The factor is usually interpreted as a shock to aggregate demand or demographics. See also Albuquerque, Eichenbaum, and Rebelo (2012).

[20] The sample period considered in this paper includes a period before the Fed was operational. The outlined model is meant to describe the more recent period where the central bank has been actively managing the short term interest rate.

The economy consists of two types of households, a representative final-good-producing firm, a continuum of intermediate-goods-producing firms and a central bank.

V.A. Households

There is a continuum of households on [0,1] where $1-\lambda$ fraction of them are forward-looking, participate in asset markets to smooth consumption and own all assets in the economy–*savers* denoted by subscript s. The *rule-of-thumb* households on $[0, \lambda]$ denoted by the subscript r do not participate in financial markets to smooth their consumption and do not have any wealth. Both types of households have an identical period utility function U which is standard:

$$U\left(C_{i,t}, N_{i,t}\right) = \frac{C_{i,t}^{1-\gamma}}{1-\gamma} - \frac{N_{i,t}^{1+\eta}}{1+\eta}, \tag{1.11}$$

where $i \in \{s, r\}$.[21] The parameter $\gamma > 0$ represents the degree of risk aversion, $\eta > 0$ is the inverse of labor supply elasticity. Savers solve the usual optimization problem:

$$\max \mathrm{E}_t \sum_{j=0}^{\infty} \beta^j U\left(C_{s,t+j}, N_{s,t+j}\right), \quad \beta \in (0,1) \tag{1.12}$$

by choosing the consumption $C_{s,t}$ and worked hours $N_{s,t}$ every period subject to the budget constraint expressed in nominal terms:

$$B_{s,t+1} + \Theta_{s,t+1} V_t \leq B_{s,t} R_t + \Theta_{s,t}\left(V_t + P_t D_t\right) + W_t N_{s,t} - P_t C_{s,t}. \tag{1.13}$$

$B_{s,t}$ is the quantity of one-period nominal bonds at the beginning of period t,[22] $\Theta_{s,t}$ denotes the holdings of savers in firms,[23] V_t is the market value of firms and D_t are real dividends. W_t denotes the nominal wage.

[21] Separable preferences are used for simplicity. The utility function given by (1.11) is not compatible with balanced growth path. To resolve this, one could introduce a non-separable utility specification. The following specification is common in the RBC literature: $U\left(C_t, N_t\right) = \frac{1}{1-\sigma}\left[\left(C_t v\left(N_t\right)\right)^{1-\sigma} - 1\right]$. The main result in this paper, namely EIS < 0, is obtained also with this specification.

[22] Nominal bonds are in the zero net supply because markets are complete and savers are homogenous. Hence, they can be replaced by a representative agent.

[23] In a model with the representative agents $\Theta = 1$. In this model, it will depend on the fraction of savers on the whole population: $\Theta = \frac{1}{1-\lambda}$.

The no-arbitrage condition implies the existence of a stochastic discount factor $M_{t,t+1}$ that prices all uncertain future cash flows:

$$V_t = \mathrm{E}_t \left[M_{t,t+1} \left(V_{t+1} + P_{t+1} D_{t+1} \right) \right]. \tag{1.14}$$

The stochastic discount factor together with the assumption of frictionless financial markets (for savers) allows to rewrite (1.13) in terms of present values:

$$\mathrm{E}_t \sum_{j=t}^{\infty} M_{t,t+j} P_j C_{s,j} \leq V_t + \mathrm{E}_t \sum_{j=t}^{\infty} M_{t,t+j} W_j N_{s,j}. \tag{1.15}$$

The remaining first order conditions for savers read:

$$\beta \frac{U_C \left(C_{s,t+1} \right)}{U_C \left(C_{s,t} \right)} = M_{t,t+1} \frac{P_{t+1}}{P_t} \tag{1.16}$$

$$\frac{1}{R_t} = \beta \mathrm{E}_t \left[\left(\frac{C_{s,t+1}}{C_{s,t}} \right)^{-\gamma} \frac{P_t}{P_{t+1}} \right] \tag{1.17}$$

$$\frac{N_{s,t}^{\eta}}{C_{s,t}^{-\gamma}} = \frac{W_t}{P_t}. \tag{1.18}$$

The optimization problem of rule-of-thumb consumers involves only the intratemporal trade-off:

$$\max_{C_{r,t}, N_{r,t}} \quad \frac{C_{r,t}^{1-\gamma}}{1-\gamma} - \frac{N_{r,t}^{1+\eta}}{1+\eta}, \tag{1.19}$$

subject to the budget restriction $C_{r,t} P_t = W_t N_{r,t}$. The first order condition for the rule-of-thumb household is:

$$\frac{N_{r,t}^{\eta}}{C_{r,t}^{-\gamma}} = \frac{W_t}{P_t}. \tag{1.20}$$

V.B. Production

The final good is produced with a CES production function with elasticity ε, which aggregates the intermediate goods indexed by k, i.e. $Y_t = \left(\int_0^1 Y_t \left(k \right)^{(\varepsilon-1)/\varepsilon} dk \right)^{\varepsilon/(\varepsilon-1)}$. Firms producing intermediate goods are characterized by linear production technology without capital:

$$Y_t(k) = A_t N_t(k) - F \text{ if } N_t(k) > F \text{ and } 0 \text{ otherwise}, \tag{1.21}$$

where F denotes the fixed costs and A_t represents the technology, where $a_t = \log A_t$ is assumed to evolve as:[24]

$$a_t = \rho_a a_{t-1} + \varepsilon_t^a, \tag{1.22}$$

with $0 \leq \rho_a \leq 1$ and $\varepsilon_t^a \sim N(0, \sigma_a^2)$. Firms producing intermediate goods face a downward-sloping demand curve: $C_t(k) = \left(\frac{P_t(k)}{P_t}\right)^{-\varepsilon} C_t$. They take wages as given and the cost minimization implies the following nominal marginal cost: $MC_t = W_t/A_t$. The total profit D_t is given by $D_t = (1 - (MC_t/P_t)\,\Delta_t)\,Y_t$, where $\Delta_t = \int_0^1 (P_t(k)/P_t)^{-\varepsilon}\,dk$ is defined as relative price dispersion. Savers, who hold shares in firms, maximize their total value by choosing price $P_t(k)$. The optimal price reads:

$$P_t^{opt}(k) = \mathrm{E}_t \sum_{j=0}^{\infty} \frac{\omega^j M_{t,t+j} P_{t+j}^{\varepsilon-1} Y_{t+j}}{\mathrm{E}_t \sum_{m=0}^{\infty} \omega^m M_{t,t+m} P_{t+m}^{\varepsilon-1} Y_{t+m}} MC_{t+j}. \tag{1.23}$$

The aggregate price index evolves as:

$$P_t^{1-\varepsilon} = (1-\omega)\left(P_t^{opt}\right)^{1-\varepsilon} + \omega P_{t-1}^{1-\varepsilon}, \tag{1.24}$$

where ω represents the fraction of firms being unable to adjust their prices, i.e. Calvo staggered pricing. The Philips curve follows from linearized versions of (1.23) and (1.24), for details see standard references, e.g. Woodford (2003b).

V.C. Equilibrium

The equilibrium requires that all markets clear. Labor markets clearing implies $N_t = \lambda N_{r,t} + (1-\lambda)\,N_{s,t}$, goods market clearing delivers $C_t = Y_t$ and $C_t = \lambda C_{r,t} + (1-\lambda)\,C_{s,t}$, and finally asset market clearing implies that $\Theta_{s,t} = \frac{1}{1-\lambda}$ for all t. Bonds are in zero net supply.

[24] The technology shock can have a unit-root, in which case the output, consumption, and real wages would need to be stochastically detrended to obtain stationary dynamics.

V.D. Key model equations

The aggregate dynamics are described by the following three equations (derivations and the discussion of steady states are in Appendix IX.E). Lower case letters below denote the log-deviation of a given variable from its steady state value.

Aggregate Euler equation

The Euler equation of savers has the usual form:

$$\mathrm{E}_t c_{s,t+1} - c_{s,t} = \frac{1}{\gamma}\left(i_t - \mathrm{E}_t \pi_{t+1}\right). \tag{1.25}$$

The output Euler equation is obtained by aggregating the consumption of savers and rule-of-thumb agents:

$$\mathrm{E}_t y_{t+1} - y_t \quad = \tfrac{1}{\delta\gamma}\left(i_t - \mathrm{E}_t \pi_{t+1}\right) + \chi\left[\mathrm{E}_t a_{t+1} - a_t\right], \tag{1.26}$$

where y_t represents the aggregate output, $\pi_t = \log\ P_t/P_{t-1}$ is the inflation and $\chi = -\frac{1}{\frac{\kappa\lambda\eta+1-\lambda}{\lambda\eta(1+\kappa)} - \frac{1}{1+\mu}}$. Estimates of $\frac{1}{\delta\gamma}$ from the aggregate intertemporal optimality condition in Table I are consistently negative. Given that savers are risk averse ($\gamma > 0$), the negative sign can arise through $\delta < 0$. The magnitude of the aggregate EIS is largely determined by δ, which is a function of structural parameters:

$$\delta = \frac{\left(\kappa\lambda\eta + 1 - \lambda\right)\left(1+\mu\right) - \lambda\eta(1+\kappa)}{\left(1-\lambda\right)\left(1+\mu\right)\left[\kappa\lambda\eta + \gamma\lambda(1+\kappa) + 1 - \lambda\right]},$$

where $\kappa \equiv \frac{1-\gamma}{\eta+\gamma}$. In case of full market participation ($\lambda = 0$) $\delta = 1$, and the standard IS curve is recovered. Importantly, δ can have either sign depending on the combination of structural parameters in the numerator.

Phillips curve

Nominal rigidities introduce a trade-off between the output gap $x_t = y_t - y_t^*$ and inflation:

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \psi\vartheta\left(y_t - y_t^*\right). \tag{1.27}$$

Natural level of output y_t^*, which depends only on the exogenous technology shock a_t, is achieved when prices are flexible, i.e. $\omega = 0$ (see Appendix IX.E for the derivation). Unlike most of the recent New-Keynesian model specifications, both the aggregate output equation and the Phillips curve are purely forward-looking.[25] In my setup, the Phillips curve is not influenced by limited participation.

Monetary policy

The central bank sets the policy rate according to the following interest rate rule:

$$i_t = \phi_i i_{t-1} + (1 - \phi_i) \left[\phi_\pi \pi_t + \phi_x \left(y_t - y_t^* \right) \right]. \tag{1.28}$$

Consistent with the empirical evidence, I assume that the central bank moves the policy rate gradually, which introduces a substantial degree of smoothing captured by ϕ_i.

V.E. Model calibration

Where possible, the choice of parameter values is guided by microeconomic evidence.

Elasticity of labor supply. Chetty, Guren, Manoli, and Weber (2011), argue that the Frisch elasticity of labor supply should be set to 0.75 which implies $\eta = 1.333$. However, the magnitude of labor supply elasticity is not uncontroversial. For this reason, I report the results for a range of values reported in the literature, $\eta \in \{1, 1.333, 2, 4\}$.

Asset market participation. I set $\lambda = 0.7$ which is motivated by the evidence from Vissing-Jorgensen (2002) who classifies 21.75 percent of households as stockholders and 31.40 percent as bondholders, based on the Consumer Expenditure Survey (CEX) for the period 1980-1996. Similarly, Guvenen (2006) argues that $\lambda = 0.7$ represents a lower bound for the fraction of non-participating households. One concern with choosing high values for λ is the fact that stock market participation of households has increased significantly in the last decades. I provide a detailed discussion of this issue in Appendix IX.D. To assess the impact of λ on my results, I vary its values between 0 and 1.

Rigidities. The elasticity of substitution among intermediate goods is set to $\varepsilon = 6$ which implies the steady-state markup value $\mu = 0.2$, (Eichenbaum and Fisher, 2007). This parameter has a minute effect on δ for any plausible markup value. The Calvo parameter of $\omega = 0.75$ implies that firms can reset prices once a year.

[25] The backward indexation of prices is usually added to the Phillips curve to make inflation more persistent and fit the data, but it has little support in micro data, e.g. (Chari, Kehoe, and McGrattan, 2009).

Risk aversion. I study parameter values ranging from 0.5 through 10. Values of γ that are substantially higher than unity are commonly used in the asset pricing literature. In contrast, Chetty (2006) obtains an upper bound for γ equal to two by exploiting the labor supply behavior.

Monetary policy. The smoothing parameter ϕ_i is set to 0.85, $\phi_\pi = 0.4$ and $\phi_x = 1.2$. In the standard New-Keynesian model without limited participation, coefficients above one for ϕ_π are necessary to ensure the determinacy (Taylor rule). However, in the case of limited participation, it is possible that $0 < \phi_\pi < 1$ leads to determinate equilibria (Gali, Lopez-Salido, and Valles, 2004; Bilbiie, 2008). In this model, π_t is a deviation of inflation from its steady state. Post-war US inflation has a stochastic trend, and Taylor rule estimates that explicitly account for it, indicate that the central bank reacted by more than one-for-one to the inflation trend, but the coefficient on the deviations from the stochastic trend is indeed less than unity. Here, I abstract from the stochastic inflation target.[26]

Technology shock. The single exogenous shock of the model is the technology shock. I set $\rho^a = 0.3$ and $\sigma_a = 0.01$.

V.F. Model solution

Equations (1.26)–(1.28) together with the exogenous technology shock determine the law of motion of the endogenous variables in the model, which is given by:

$$z_t = \Gamma z_{t-1} + \Xi a_t, \tag{1.29}$$

where $z_t = [\pi_t, y_t, i_t, y_t^*]'$. More details about the model solution are provided in Appendix IX.E. Table VI collects all the parameters. Numerically solving the model for these parameter values gives the following dynamics:

$$\Gamma = \begin{bmatrix} 0 & 0 & 0.9570 & 0 \\ 0 & 0 & 0.2224 & 0 \\ 0 & 0 & 0.9475 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Xi = \begin{bmatrix} 0.3335 \\ 1.2423 \\ 0.1087 \\ 0.7496 \end{bmatrix}. \tag{1.30}$$

[26]See Cogley and Sbordone (2008) and Coibion and Gorodnichenko (2011b) for the analysis of the New Keynesian models with inflation trend.

Table VI: Model parameters

This table collects the parameter values used to calibrate the model. All parameters refer to quarterly frequency.

Parameter		Value
Preference parameters		
Discount factor	β	0.99
Risk aversion	γ	2
Labor elasticity parameter	η	1.33
Price rigidities		
Steady state markup	μ	0.2
Calvo parameter	ω	0.75
Limited participation		
Degree of limited participation	λ	0.7
Interest rate rule		
Interest rate smoothing	ϕ_i	0.85
Inflation sensitivity	ϕ_π	0.4
Output gap sensitivity	ϕ_x	1.2
Exogenous shocks		
Persistence of technology shock	ρ_a	0.3
Volatility of technology shock	σ_a	0.01

The single most important observation about (1.30) is that the dynamics of endogenous variables are persistent. This is surprising given that the Phillips curve and the aggregate Euler equation are purely forward-looking. A full participation model does not generate such a persistence. Thus, limited market participation, could replace habits and backward indexation, both of which are typically relied upon but lack strong support in the data.

VI. Model discussion

This section inspects the key mechanism in the model that can switch the sign of the aggregate EIS. Subsequently, it discusses the link between the aggregate consumption and the consumption of savers. Finally, it shows how the negative sign of the EIS helps reconcile the consumption-based models with the data.

VI.A. Justifying $\tilde{r}_t$ as a proxy for the real rate

I simulate the model outlined above and construct a term structure of yields implied by this model. The details of the New-Keynesian term structure model and of the simulation exercise are provided in Appendix IX.E. The simulation shows that $\tilde{r}_t$, which is the negative of the slope of the simulated term structure, closely tracks the real rate constructed as $i_t - \mathrm{E}_t\pi_{t+1}$, where an AR(1) model for realized inflation is used to compute $\mathrm{E}_t\pi_{t+1}$. Median correlation is around 0.7, see Figure 1.11 in Appendix IX.E. As can be seen from the law of motion of endogenous variables given by (1.30), inflation is more persistent than the output. Therefore, the slope captures the variation in the real rate.

VI.B. What makes δ negative?

Figure 1.6 plots the values of the aggregate EIS given by $\frac{1}{\delta\gamma}$ for a range of structural parameters γ, λ, and η together with its estimate from the total consumption data in the period 1875-2009 (Panel A of Table I). Parameter values for the degree of limited asset market participation λ, the risk aversion parameter γ and the labor supply elasticity η, motivated above are consistent with the estimate of $\frac{1}{\delta\gamma}$.

Conditional on a moderate degree of risk aversion $\gamma = 2$, Figure 1.7 displays two key ingredients that drive the sign of δ: (i) inelastic labor supply as measured by η and, (ii) limited market participation. δ changes its sign at the threshold value of market participation defined as $\hat{\lambda}$. Appendix IX.E derives $\hat{\lambda}$ in terms of structural parameters.[27] For the whole range of labor supply elasticities considered, the sign of δ switches for non-participation levels between 0.25 and 0.55.

The intuition for the negative aggregate EIS is as follows. A positive exogenous technology shock increases the natural level of output y_t^* and thus changes the output gap x_t. Depending on the configuration of the parameter values, in particular those in the monetary policy reaction function, the increase in y_t^* translates through the monetary policy reaction into a positive or negative shock to the real interest rate. In the full participation case, a positive shock to the real interest rate reduces the demand today as it induces the agents to seize better investment opportunities and substitute today's consumption into the future. In the limited participation model, only savers react to the real rate shock and substitute consumption intertemporally. Intertemporal choices of savers shift the labor supply curve (derived in Appendix IX.E):

[27] The aggregate EIS has a discontinuity point at $\hat{\lambda}$. For this reason, the negative sign of the EIS cannot be achieved by introducing agents with different non-zero EIS (Guvenen, 2006). To obtain the negative sign, a fraction of households needs to be excluded from all asset markets.

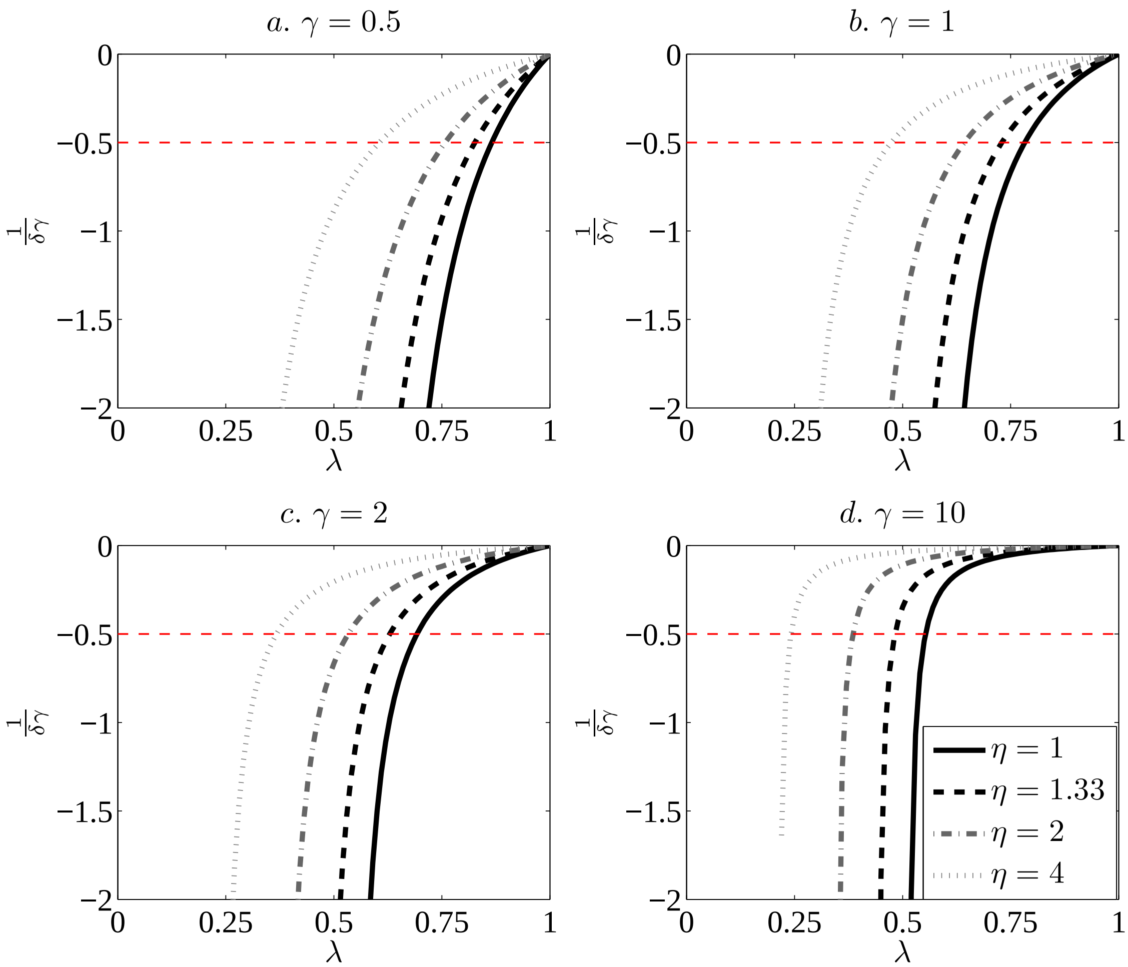


Figure 1.6: Parameterizations of the output Euler equation.

Panels $a-d$ plot the parameter values for $\frac{1}{\delta\gamma}$ in the aggregate output Euler equation given by (1.26). Three structural parameters are varied. First, the coefficient of relative risk aversion γ takes values between 0.5 and 10. Each panel represents different values of γ. Second, the Frisch elasticity of labor supply is varied between 0.25 and 1 which implies parameter values for η between 1 and 4. Third, the share of rule-of-thumb consumers λ varies between 0 and 1 (x-axis). The red dashed line represents the estimate of $\frac{1}{\delta\gamma}$ based on the annual data for total consumption and interest rates in the period 1875-2009, see the first column of Panel A in Table I.

$$w_t = \frac{\eta}{\eta\lambda\kappa + (1-\lambda)} n_t + \frac{(1-\lambda)\gamma}{\eta\lambda\kappa + (1-\lambda)} c_{s,t}. \tag{1.31}$$

For the parameter values given in Table VI, $\frac{(1-\lambda)\gamma}{\eta\lambda\kappa+(1-\lambda)} > 0$, i.e. a positive shock to the real interest rate reduces $c_{s,t}$ and thus the real wage w_t. The lower real wage further reduces the output because rule-of-thumb agents consume their wages. So far, this is the strengthened standard logic by which a positive shock to the real interest rate induces a fall in output.

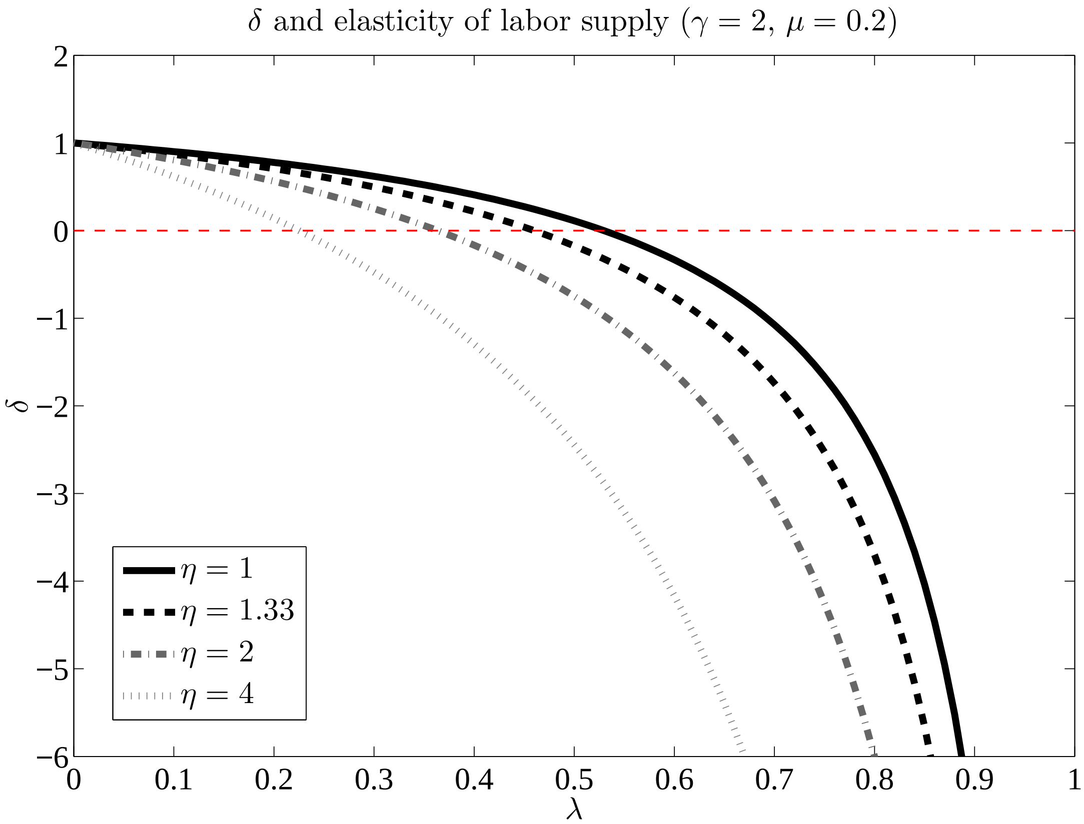


Figure 1.7: Parameterization of δ.

The figure shows the dependence of δ on the elasticity of labor supply parameter η and the asset market participation parameter λ. Expression for δ is given by (1.27). Note that η is the inverse of the Frisch elasticity of labor supply. Hence, the elasticity of labor supply is varied between 0.25 and 1. The risk aversion parameter is $\gamma = 2$ and the steady-state markup is $\mu = 0.2$.

To invert this logic, one needs to consider the dividend income of savers. As can be seen directly from the budget constraint, the consumption of savers has two components, wages and dividends:

$$c_{s,t} = w_t + n_{s,t} + \frac{1}{1-\lambda} d_t, \tag{1.32}$$

where real profits d_t, accruing only to savers, are given by:

$$d_t = \frac{\mu}{1+\mu} y_t - w_t + a_t. \tag{1.33}$$

If the fall in the aggregate output y_t brought about by the interest rate increase is paired with a correspondingly larger decrease in wages, it leads to a positive income effect for savers (equation (1.33)). The positive income effect generates additional labor demand and the real wage increases. In equilibrium, both the aggregate output and real wages increase. Hence, a higher real interest rate coincides with an increase in aggregate output induced by the positive income effect for savers. This mechanism implies that firm markups are counter-cyclical which is consistent with the empirical evidence (Rotemberg and Woodford, 1999).

VI.C. Consumption of savers and rule-of-thumb consumers

The negative δ implies that the consumption growth of rule-of-thumb consumers is negatively correlated with the variation in the real interest rate. This link arises as a result of intertemporal choices of savers and does not imply that rule-of-thumb consumers have a negative EIS. Consistent with this prediction, using UK household-level data, Attanasio, Banks, and Tanner (2002) report a negative relationship for non-shareholders (Panel C of Table II in their paper).

The model implies that the consumption of savers loads negatively on the aggregate consumption and positively on technology shocks, which can be seen in equation (1.34) below:

$$c_{s,t} = \delta y_t + \zeta a_t, \tag{1.34}$$

where $\zeta > 0$ for $\lambda > 0$ and is increasing in λ:

$$\zeta = \frac{\lambda\eta(1+\kappa)}{(1-\lambda)\left[\kappa\lambda\eta - \lambda + \gamma\lambda(1+\kappa) + 1\right]},$$

(derivation of (1.34) and ζ is in Appendix IX.E). Depending on the magnitude of δ and ζ, the consumption growth of savers is more volatile than the aggregate consumption growth. For $\delta < 0$ and moderate risk aversion parameter values, i.e. $\gamma \ll 10$, we have that $\zeta > 1$ as illustrated in Panels *a*–*d* of Figure 1.8. With larger λ, a smaller fraction of savers holds all stocks in the economy and is thus more exposed to technology shocks.[28] Equation (1.34) does not imply that the consumption of savers is negatively correlated with the aggregate consumption in the data. The technology shock a_t is positively correlated with y_t and can induce a positive correlation in the data despite negative δ. Equation (1.34) also helps

[28] Setting $\lambda = 0$ implies $\delta = 1$ and $\zeta = 0$, which restores the representative agent setup.

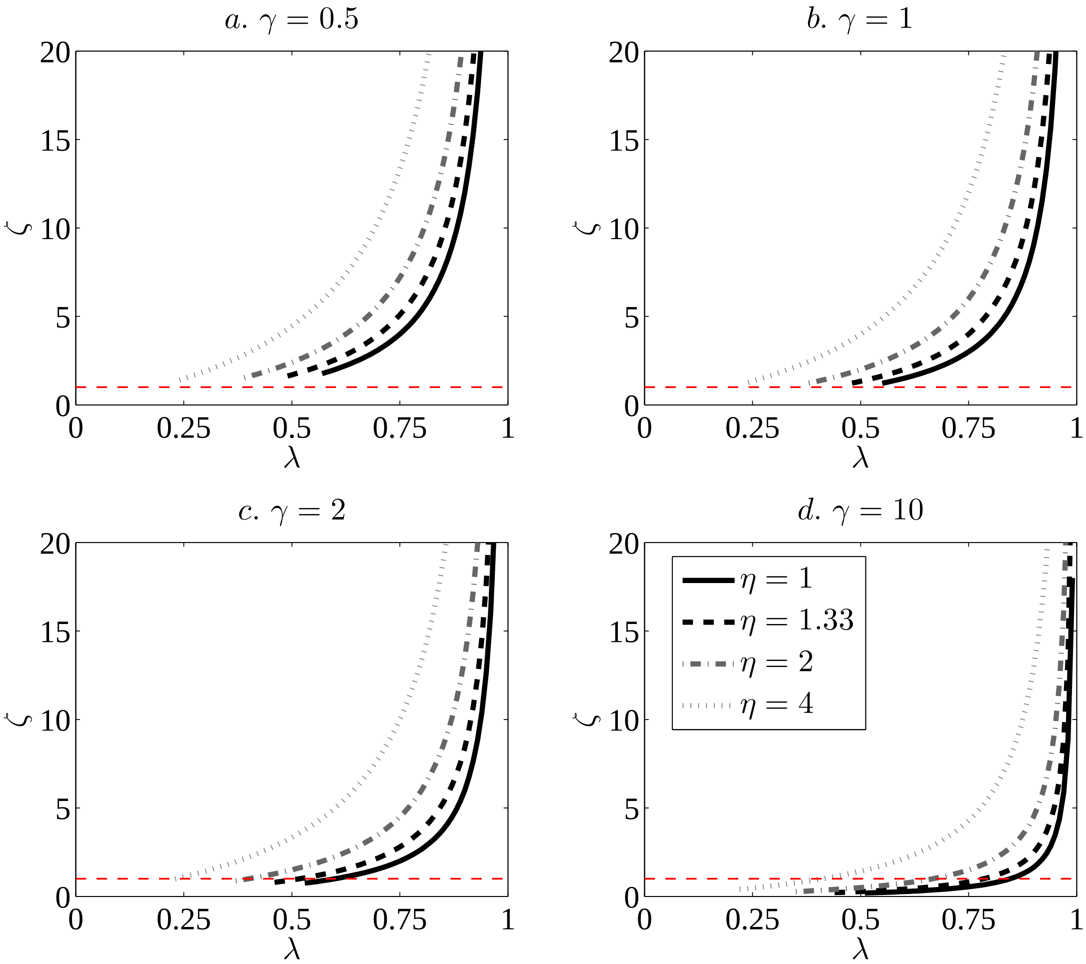


Figure 1.8: Exposure of consumption of savers to the technology shocks.

Panels $a - d$ plot the parameter values for ζ in the equation linking the consumption of savers to the total output given by (1.34). Three structural parameters are varied. First, the coefficient of relative risk aversion γ takes values between 0.5 and 10. Each panel represents different values of γ. Second, the Frisch elasticity of labor supply is varied between 0.25 and 1 which implies parameter values for η between 1 and 4. Third, the share of rule-of-thumb consumers λ varies between 0 and 1 (x-axis). All plotted values of ζ are obtained for $\delta < 0$.

explain why the aggregate consumption appears too smooth when compared to substantial volatility of returns on risky assets. The model suggests that it is not driven by efficient risk sharing but rather by intertemporal choices of savers which induce the negative correlation of the consumption of savers with the total consumption.

VI.D. Why does the standard consumption Euler equation fail?

The Euler equation for aggregate consumption fails to replicate the variation and the level of interest rates in the data (Weil, 1989). Recently, Cumby and Diba (2007) show that a broad range of preference specifications in consumption-based models imply a real short rate that is either negatively correlated with the observed ex-post real interest rate or it is extremely volatile and uncorrelated with the observed one.

The model outlined in this paper offers one way to reconcile the mismatch between the consumption-based models and observed short-term yield dynamics. The log of nominal pricing kernel implied by the model reads:

$$m_{t,t+1} = \log(\beta) - \gamma \Delta c_{s,t+1} - \pi_{t+1}, \tag{1.35}$$

where $\Delta c_{s,t+1} = \log C_{s,t+1} - \log C_{s,t}$ denotes the consumption growth of savers. Using the differenced version of (1.34) and the market clearing condition $y_t = c_t$, rewrite (1.35) as:

$$m_{t,t+1} = \log(\beta) - \gamma\delta\Delta c_{t+1} - \gamma\zeta\Delta a_{t+1} - \pi_{t+1}. \tag{1.36}$$

The comparison of (1.35) with (1.36) provides the intuition for the mismatch between the observed short term interest rates and the pricing kernel. For $\delta < 0$, which is empirically the case, the aggregate consumption enters the pricing kernel with the opposite sign compared to the consumption of savers who are marginal bond investors. The presence of rule-of-thumb consumers introduces a wedge between the aggregate Euler equation and the Euler equation of bond market participants (savers). Given that the consumption of savers is not directly observable, when equating the aggregate consumption Euler equation with observed interest rates, one needs to use $\frac{1}{\delta\gamma}$.

The literature has proposed solutions to the empirical failure of the consumption Euler equation that are largely risk-based and/or depend on the specifics of monetary policy (Gallmeyer, Hollifield, Palomino, and Zin, 2007; Atkeson and Kehoe, 2008; Reynard and Schabert, 2012).[29] However, the negative relationship between real interest rates and consumption growth appears to be structural, hence exists irrespective of monetary policy.

[29]Models with time-varying risk premia add the stochastic second-order terms into the Euler equation while keeping the first-order relationship unchanged. This is done either through time-varying risk aversion (e.g. habits) or stochastic volatility (e.g. long-run risk models).

Indeed, the results in Table I indicate that the relationship exists across periods marked by very different monetary policy regimes.

VII. Is the negative EIS consistent with long-run risk models?

Empirical results in previous sections imply that the expected consumption growth varies in a persistent way. This section shows that the persistent component of expected consumption growth rate extracted from the Euler equation with the negative EIS has asset pricing implications that are consistent with models featuring long-run cash flow risk (Bansal and Yaron, 2004; Hansen, Heaton, and Li, 2008).

In the standard long-run risk model (Bansal and Yaron, 2004), both the consumption and dividend growth, g_t and $g_{d,t}$ contain the persistent component x_t and their innovations have stochastic volatility σ_t^2. Further consider Epstein-Zin preferences with γ and σ denoting the risk aversion and the EIS, respectively. The essence of the model are two equations given by (1.37)–(1.38). Both, the short term real rate r_t and the log price-dividend ratio $z_{m,t}$ are affine functions of x_t and σ_t^2:

$$r_t = A_{0,f} + A_{1,f}x_t + A_{2,f}\sigma_t^2, \tag{1.37}$$

$$z_{m,t} = A_{0,m} + A_{1,m}x_t + A_{2,m}\sigma_t^2, \tag{1.38}$$

where $A_{1,f} = \frac{1}{\sigma}$.[30] Hence, the expected consumption growth x_t has a first-order effect on the real rate r_t. σ_t^2 represents the fluctuations in risk and is the key determinant of risky asset valuations represented by $z_{m,t}$. The long-run risk literature restricts its attention to positive values of preference parameters. Therefore, both the EIS and the risk aversion parameter have to be greater than unity to get plausible asset pricing implications, i.e. price-dividend ratio increasing in x_t and decreasing in σ_t^2. However, it can readily be seen from $\theta = \frac{1-\gamma}{1-\frac{1}{\sigma}}$ that the negative value of EIS also implies $\theta < 0$ and thus preserves the desirable asset pricing properties of long-run risk models.

To empirically evaluate the asset pricing properties of x_t, I first extract the persistent component of consumption growth. From (1.37), we have:

[30] The solution coefficients $A_1, A_2, A_{0,m}, A_{1,m}, A_{2,m}, A_{0,f}, A_{1,f}, A_{2,f}$ are defined in Bansal and Yaron (2004) and are not repeated here for brevity.

$$x_t = -\sigma A_{0,f} + \sigma r_t - \sigma A_{2,f}\sigma_t^2. \tag{1.39}$$

The affine relationship given in (1.39) is implemented via OLS:

$$\Delta c_{t+1} = \delta_0 + \underbrace{\delta_1}_{-0.81\ [-5.23]} \tilde{r}_t + \underbrace{\delta_2}_{0.005\ [1.32]} \hat{\sigma}_t^2 + \varepsilon_{t+1} \quad \bar{R}^2 = 0.11, \tag{1.40}$$

where $\hat{\sigma}_t^2$ is an empirical proxy for the consumption volatility: $\hat{\sigma}_t^2 = \log \sum_{i=1}^{5} |\eta_{t-i}|$ with η_t being the innovation in consumption growth obtained from an AR(1) estimated on Δc_t recursively using the window size of 30 years. The regression results show that the volatility of consumption growth is unrelated to future consumption growth. Contrary to the argument that stochastic volatility induces a downward bias in the EIS estimates from the linearized Euler equation, the estimated δ_1 in (1.40) shows that this is not the case. The extracted persistent component denoted by $\hat{x}_t$, is the fitted value from a restricted version of (1.40) with $\delta_2 = 0$. This is justified given the insignificant estimate of δ_2 in the unrestricted version. Figure 1.9 plots $\hat{x}_t$ together with the realized consumption growth. Consistent with the basic intuition, the expected consumption growth explains a fairly small portion ($\bar{R}^2 = 0.11$) of the overall variation in consumption growth.

One of the key implications of the long-run risk model is that x_t can be predicted by the price-dividend ratio which is based on the relationship given by (1.38). However, establishing this link empirically has been a challenge. I use $\hat{x}_t$ to estimate (1.38):

$$z_{m,t} = \gamma_0 + \underbrace{\gamma_1}_{7.97\ [2.63]} \hat{x}_t + \underbrace{\gamma_2}_{-0.48\ [-6.02]} \hat{\sigma}_t^2 + \epsilon_t \quad \bar{R}^2 = 0.57. \tag{1.41}$$

In both regressions, (1.40) and (1.41), Newey-West-adjusted t-statistics with three lags are reported in parentheses. Due to data limitations, regression (1.41) is estimated on the sample 1926-2009 using annual data.[31] Importantly, both $\hat{x}_t$ and $\hat{\sigma}_t^2$ are statistically significant and have the expected signs. Higher expected consumption growth rate increases the price-dividend ratio while an increase in consumption volatility depresses it. This result provides supportive evidence for the validity of $\hat{x}_t$ as a persistent component of consumption

[31] In constructing the price-dividend ratio $z_{m,t}$, I follow the literature and infer the dividend payments from monthly returns on the value-weighted portfolio of all NYSE, Amex, and NASDAQ stocks. Monthly dividend distributions are reinvested in the three-month T-bill. These data are available starting in 1926 which determines the start of the sample for (1.41). Correspondingly, $\hat{x}_t$ used in (1.41) is constructed using the data from the 1926-2009 sample.

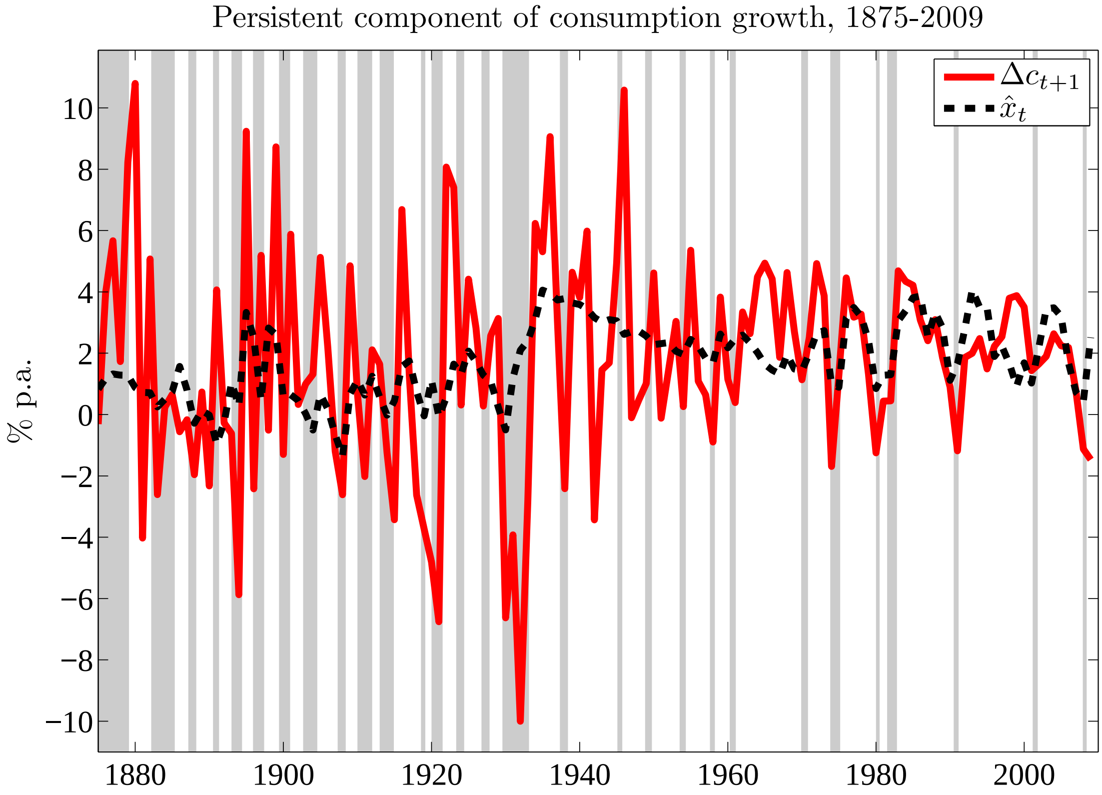


Figure 1.9: Time-varying expected consumption growth, 1875-2009

The figure plots the annual consumption growth together with the estimated persistent component $\hat{x}_t$, which is obtained as a fitted value from the regression of Δc_{t+1} on $\tilde{r}_t$. Grey areas represent the recessions. The data are annual and the sample period is 1875 through 2009.

growth.[32] The key takeaway from this exercise is that additional moments in the form of equity valuation ratios are consistent with the negative value of the EIS. However, this is not to say that working with the negative EIS is appropriate. Optimally, one should use the EIS close to unity but apply it to the consumption data aggregated only over households actively participating in asset markets as has been done in Malloy, Moskowitz, and Vissing-Jorgensen (2009). In other words, the discrepancy between the aggregate EIS estimates and the micro evidence is largely a result of aggregation.

[32]Note that $\hat{x}_t$ can be extracted in real-time as $\tilde{r}_t$ is constructed from asset prices. Estimating (1.40) recursively gives a stable and significant estimate of δ_1.

VIII. Conclusions

This paper shows that the aggregate elasticity of intertemporal substitution is negative around -0.5. This empirical finding might seem counterintuitive at first but it is sensible if one considers that a substantial fraction of aggregate consumption is attributed to agents whose financial wealth is negligible. These agents have little incentive to participate in asset markets and largely consume their wage income. This single friction is responsible for the negative sign of the aggregate EIS and helps explain several empirical puzzles in the asset pricing and monetary policy literature.

IX. Appendix

IX.A. A two-component model for inflation

The US inflation can be well described by a two-component model with stochastic volatility (UC-SV), (Stock and Watson, 2007):

$$\begin{aligned}
\pi_t &= \tau_t + \eta_t \quad \text{with } \eta_t = \sigma_\eta \zeta_{\eta,t} \\
\tau_t &= \tau_{t-1} + \varepsilon_t \quad \text{with } \varepsilon_t = \sigma_\varepsilon \zeta_{\varepsilon,t} \\
\ln \sigma^2_{\eta,t} &= \ln \sigma^2_{\eta,t-1} + \nu_{\eta,t} \\
\ln \sigma^2_{\varepsilon,t} &= \ln \sigma^2_{\varepsilon,t-1} + \nu_{\varepsilon,t},
\end{aligned}$$

where $\zeta_t = (\zeta_{\eta,t}, \zeta_{\varepsilon,t})$ is i.i.d. $N(0, I_2)$, $\nu_{\eta,t} \sim N(0, \gamma_1)$ and $\nu_{\varepsilon,t} \sim N(0, \gamma_2)$. All four innovations are mutually independent.

The simulation exercise replicates the linearized Euler equation given by (1.8), falsely assuming constant volatility. In the first step, I construct the nominal interest rate by simulating the real rate and inflation. I simulate realized inflation from the UC-SV model above and set $\gamma_1 = 0.2$ and $\gamma_2 = 0.5$ to reflect the higher volatility fluctuations in the transitory component of inflation, observed in the early part of the 1875-2009 sample. I set T=135 to be consistent with the sample size used in this paper. I simulate the real rate as an AR(1) process, assuming the following: $\phi_r = 0.75$, $\sigma_r = 1$ and $\bar{r} = 1.9$, where the data are annual percentages. For consumption growth, I assume $\sigma_c = 3.2$ and EIS=1. The ex-ante real rate is obtained by linearly projecting the ex-post real rate observed at $t + 1$ on past inflation and the nominal interest rate. I simulate the model 1000 times and report median values.

IX.B. Data

This section details the data series used to estimate the consumption and output Euler equations.

Yield data. I study US yields in the period January 1875 through December 2009 unless otherwise stated. The raw yield data are monthly and are obtained from the Global Financial Data (GFD) database. I consider the ten-year constant maturity yield (series IGUSA10D) as long term yield. To construct the short term yield for the whole sample period, I combine the three-month AA non-financial commercial paper yield (series IPUSAC3D) (1875-1933) with the three-month Treasury bill (1934-2009). The reason for using the three-month commercial paper in the earlier part of the sample is the data availability. The ten-year constant maturity yield is interpolated from available long term non-inflation-indexed Treasury securities. GFD itself combines various data sources such as the National Monetary Statistics from the Federal Reserve to obtain the ten-year CMT yield. $\tilde{r}_t$ is an average of monthly observations of $-\text{slope}_t$ within each year.

Consumption and output data (1875-2009). The consumption and output data are constructed by Robert Barro and Jose Ursua and downloaded from their website.[33] These data contain the total real per capita consumption/output index without further split into durable and non-durable consumption. Barro and Ursua compile the consumption data from various sources: For the period 1869-1899, the data are from Rhode (2002). For the 1900-1928 period, consumption data are taken from Lebergott (1996) and the period 1929-2009 is covered by data from the Bureau of Economic Analysis. The output data are sourced from Balke and Gordon (1989) for the period 1869-1928 and from BEA for the 1928-2009 period.

Consumption data (1929-2009). The consumption data for this period are obtained from the Bureau of Economic Analysis (US Dept. of Commerce), Tables 2.3.4 and 2.3.5. Non-durable goods and services as well as durables are deflated by their corresponding price indexes (2005=100). Population data are obtained from the Bureau of Labor Statistics for the period 1948-2009. In the period 1929-1947, the population over 16 years is obtained from the Population Estimates Program of the U.S. Census Bureau[34] by dividing the whole population by 1.5. This ratio of population over 16 years to the total population has been stable in the 1950s.

Output and non-Treasury interest rate data (1869-1983). The real GNP data for the period 1869-1983 are from Table I, Appendix B in Balke and Gordon (1986). The data in Balke and Gordon (1986) are assembled from various sources, see Notes on Section 1 of their paper for more details. The proxy for the real rate $\tilde{r}_t$ is computed as a difference between the commercial paper rate and yield on corporate bonds. The commercial paper has a maturity of six-months in most of the sample. Yields on corporate bonds are Baa-rated (Moodys) corporate bond yields (1919-1983) and railroad bond yields (1869-1918).

International data

Consumption and output data (1950-2009). Annual consumption and output data for Canada, Germany, Netherlands, Switzerland and the United Kingdom are obtained from the Barro and Ursua's data set.

[33] The dataset is available at http://rbarro.com/data-sets/.

[34] http://www.census.gov/popest/data/national/totals/pre-1980/tables/popclockest.txt.

Yield data. All yield data are from the GFD database, see their data descriptions for the original sources. For all countries, short-term nominal yield is represented by a three-month Treasury bill. The only exception is Switzerland in the period 1950-1979 where the commercial paper rate is used instead. The long-term yield is represented by a ten-year Treasury bond yield for Canada, Germany and Netherlands. For the UK, I use the 20-year government bond and the conferedate long-term bond for Switzerland. $\tilde{r}_t$ is a difference between the three-month Treasury bill yield and the long-term bond yield averaged over monthly observations in a given year.

IX.C. EIS estimation: non-linear GMM and capital taxes

There are at least two reasons for estimating the log-linearized version of the Euler equation as opposed to estimating equation (1.6) via non-linear GMM (Hansen and Singleton, 1982). First, it is likely that consumption is measured with error which, if present, causes a more severe bias in the non-linear estimation. This is especially relevant for the consumption data in the period before 1947. Romer (1986) shows that macro data before World War II are less accurate and are constructed using a different methodology than the post-World War II data. Second, $\tilde{r}_t$ is not informative about the unconditional mean of the real rate. Given that I want to avoid the use of the ex-post real interest rate in the estimation, the non-linear estimation is not feasible.

I estimate the aggregate EIS from the pre-tax data. Applying the capital income tax rate to $\tilde{r}_t$ directly is not appropriate because this measure does not represent the real interest income but rather its variation. Additionally, Mishkin (1981) argues that the effective tax rate varies substantially across households which makes it difficult to know the appropriate tax rate at macro level. Therefore, I do not adjust $\tilde{r}_t$ for capital income taxes which might bias the EIS estimates toward zero. Hence, my estimates of the EIS represent a lower bound on its magnitude.[35]

IX.D. Empirical evidence on asset market participation

Wolff (2004) reports that the fraction of households with direct holdings of stocks increased from 13.10 percent in 1989 to 21.30 percent in 2001. Moreover, the share of stock-owning households including indirect holdings through mutual funds or pension accounts rose even more from 31.70 to 51.90 percent in the same period.

[35] For illustration, the average marginal tax rate for capital income was slightly below 5% before World-War II, see Table 2 in Barro and Sahasakul (1983). Since the mid-1940s, the rate has been increasing to a level around 30%.

However, these numbers might misrepresent the degree of asset market participation for the purpose of this model. In the model, λ represents the fraction of households that actively participate in asset markets to smooth consumption which requires substantial liquid wealth. According to Wolff (2004), only 35 percent of households owned USD 10 000 (in 1995 dollars) or more of stocks in 2001. Using the Survey of Consumer Finances (SCF), Poterba and Samwick (1995) report that 24.60 and 29.30 percent of households had direct and indirect stock holdings of more than USD 2 000 (in 1992 dollars) in 1983 and 1992, respectively.

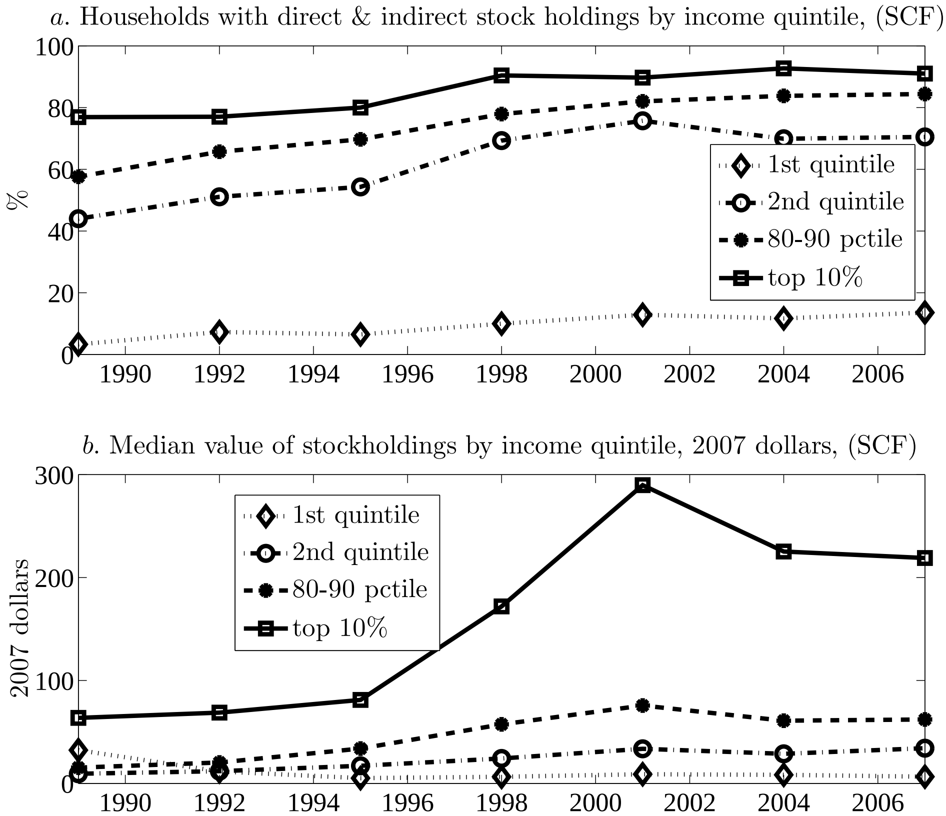


Figure 1.10: Stock market participation and holdings by income, 1989-2007

Panel *a* plots the percentiles of households with direct or indirect stock holdings split by their income. Two lower quintiles and two upper deciles are displayed for the period 1989-2007. Panel *b* plots the median value of stock holdings of households by income. The values are in 2007 dollars and thus directly comparable across time. The data are obtained from the Survey of Consumer Finance.

I use the SCF data to show that the financial wealth of low income households has not risen in real terms in the last two decades. Panel *a* of Figure 1.10 reports the fraction of stockholders by income quintiles for the period 1989-2007, and Panel *b* reports the corresponding median values of stock holdings adjusted for inflation. The important observation is that while the degree of participation has increased, the value of stock holdings for the lowest income quintile has remained virtually unchanged in the last two decades. In line with this evidence, Carroll (2000) reports that according to the SCF, the lowest 66 percentiles of the US population by wealth have liquid assets of only 10 percent of their annual wage income. Overall, these numbers indicate that $\lambda = 0.7$ is justified. The likely reason for the limited asset market participation are various forms of costs attached to it. Section 5 in Vissing-Jorgensen (2004) provides a thorough discussion of participation costs and their ability to explain the non-participation of a substantial portion of households.[36]

IX.E. Model

This section provides details on some derivations of the model outlined in Section II.

Steady state

The steady state risk-less interest rate is $R = \frac{1}{\beta}$, the net mark-up is $\mu = \frac{1}{\varepsilon - 1}$ and define $F_Y = F/Y$. Fixed costs are introduced for convenience. Following Bilbiie (2008), I set the fixed costs share $F_Y = \mu$ which implies that profits are zero in steady-state. The share of the real wage on the total output is: $WN/PY = (1 + F_Y)/(1 + \mu)$, share of profit is $D/Y = (\mu - F)/(1 + \mu)$. Assume that both savers and rule-of-thumb consumers have the same hours worked in the steady state: $N_s = N_r = N$. It follows:

$$\frac{C_s}{Y} = \frac{1 + F_Y}{1 + \mu} + \Theta \frac{\mu - F_Y}{1 + \mu}$$

using the assumption $F_Y = \mu$, we have:

$$\frac{C_s}{Y} = 1$$

$$\frac{C_r}{Y} = \frac{1 + F_Y}{1 + \mu} = 1,$$

Hence, both types of agents have the same consumption in the steady state: $C_{s,t} = C_{r,t} = \left[\frac{N^\eta P}{W}\right]^{-1/\gamma}$.

Aggregate Euler equation

The aggregate Euler equation is derived by manipulating the the equilibrium conditions and budget constraints of both types of agents. First, I provide the derivation of the threshold level of

[36] Vissing-Jorgensen (2004) estimates a median per-period cost of participation of \$350 (real 1982-84 dollars).

market participation $\hat{\lambda}$ which is the discontinuity point at where the sign of the output Euler equation switches. Combining the budget constraint of rule-of-thumb consumers and their first order condition we get:

$$\begin{aligned} N_{r,t}^{\eta} &= \left(\frac{W_t}{P_t}N_{r,t}\right)^{-\gamma}\frac{W_t}{P_t} \\ N_{r,t}^{\eta+\gamma} &= \left(\frac{W_t}{P_t}\right)^{1-\gamma}. \end{aligned}$$

Taking logs and rearranging yields:

$$n_{r,t} = \frac{1-\gamma}{\eta+\gamma}w_t. \tag{1.42}$$

To simplify the notation, define $\kappa \equiv \frac{1-\gamma}{\eta+\gamma}$. From the budget constraint of rule-of-thumb consumers we have: $c_{r,t} = (1+\kappa)w_t$. Total consumption c_t and hours worked n_t are defined as:

$$c_t = \lambda c_{r,t} + (1-\lambda)c_{s,t} \tag{1.43}$$

$$n_t = \lambda n_{r,t} + (1-\lambda)n_{s,t}. \tag{1.44}$$

Combining (1.42) and (1.44) with the labor supply of savers, we obtain the expression for the wage:

$$w_t = \frac{(1-\lambda)\gamma}{\eta\lambda\kappa+(1-\lambda)}c_{s,t} + \frac{\eta}{\eta\lambda\kappa+(1-\lambda)}n_t. \tag{1.45}$$

Plugging the consumption of rule-of-thumb consumers into (1.43) and using (1.45) yields:

$$c_t = \frac{[(1-\lambda)\left[\gamma\lambda\left(1+\kappa\right)+\left(\eta\lambda\kappa+\left(1-\lambda\right)\right)\right]]}{\eta\lambda\kappa+(1-\lambda)}c_{s,t} + \frac{\lambda\eta(1+\kappa)}{\eta\lambda\kappa+(1-\lambda)}n_t. \tag{1.46}$$

Expressing (1.46) in terms of output gives:

$$c_t = \frac{[(1-\lambda)\left[\gamma\lambda\left(1+\kappa\right)+\left(\eta\lambda\kappa+\left(1-\lambda\right)\right)\right]]}{\eta\lambda\kappa+(1-\lambda)}c_{s,t} + \frac{\lambda\eta(1+\kappa)}{\eta\lambda\kappa+(1-\lambda)}\frac{1}{1+\mu}y_t - \frac{\lambda\eta(1+\kappa)}{\eta\lambda\kappa+(1-\lambda)}a_t. \tag{1.47}$$

The differenced version of equation (1.47) is similar to the regression used in Campbell and Mankiw (1989) to estimate the fraction of rule-of-thumb households. The co-movement of consumption and output arises endogenously in the model. However, $\frac{\lambda\eta(1+\kappa)}{\eta\lambda\kappa+(1-\lambda)}\frac{1}{1+\mu}$ is not linear in λ, therefore linear regressions yield biased estimates of the fraction of rule-of-thumb households. The output Euler equation changes its sign when $\frac{\partial c_t}{\partial y_t} > 1$, therefore the threshold reads:

$$
\begin{aligned}
1 &= \frac{\lambda\eta(1+\kappa)}{\eta\lambda\kappa + (1-\lambda)} \frac{1}{1+\mu} \\
\lambda\eta(1+\kappa) &= (1+\mu)\eta\lambda\kappa + (1-\lambda)(1+\mu) \\
\hat{\lambda} &= \frac{1+\mu}{1+\mu-(1+\mu)\eta\kappa + \eta(1+\kappa)} \\
&= \frac{1}{1+\frac{\eta(1-\kappa\mu)}{1+\mu}}.
\end{aligned}
$$

The risk aversion γ influences the threshold value $\hat{\lambda}$ in a non-linear way through κ. Rearrange (1.47) and impose the clearing condition $c_t = y_t$ to obtain:

$$
c_{s,t} = \underbrace{\frac{(\eta\lambda\kappa + 1 - \lambda)(1+\mu) - \lambda\eta(1+\kappa)}{(1-\lambda)(1+\mu)\left[\kappa\lambda\eta + \gamma\lambda(1+\kappa) + 1 - \lambda\right]}}_{\delta} y_t + \underbrace{\frac{\lambda\eta(1+\kappa)}{(1-\lambda)\left[\kappa\lambda\eta - \lambda + \gamma\lambda(1+\kappa) + 1\right]}}_{\zeta} a_t. \tag{1.48}
$$

The Euler equation of savers reads:

$$
\mathrm{E}_t c_{s,t+1} - c_{s,t} = \frac{1}{\gamma}\left(i_t - \mathrm{E}_t\pi_{t+1}\right), \tag{1.49}
$$

substituting (1.48) into the Euler equation of savers obtains the aggregate Euler equation:

$$
\mathrm{E}_t y_{t+1} - y_t = \frac{1}{\delta\gamma}\left(i_t - \mathrm{E}_t\pi_{t+1}\right) - \frac{1}{\frac{\kappa\lambda\eta+1-\lambda}{\lambda\eta(1+\kappa)} - \frac{1}{1+\mu}}\left[\mathrm{E}_t a_{t+1} - a_t\right] \tag{1.50}
$$

$$
\mathrm{E}_t y_{t+1} - y_t = \frac{1}{\delta\gamma}\left(i_t - \mathrm{E}_t\pi_{t+1}\right) + \chi\left[\mathrm{E}_t a_{t+1} - a_t\right], \tag{1.51}
$$

where $\chi \equiv -\frac{1}{\frac{\kappa\lambda\eta+1-\lambda}{\lambda\eta(1+\kappa)} - \frac{1}{1+\mu}}$.

Natural real rate

I start by deriving the relationship between the real wage and total output. Combining (1.48) with the production function given by $y_t = (1+\mu)a_t + (1+\mu)n_t$ and the labor supply of savers given by $\eta n_{s,t} = w_t - \gamma c_{s,t}$ and rearranging obtains:

$$
w_t = \underbrace{\frac{\eta + (1+\mu)(1-\lambda)\gamma\delta}{(1+\mu)(1-\lambda+\lambda\eta\kappa)}}_{\vartheta} y_t + \underbrace{\frac{(1-\lambda)\gamma\zeta - \eta}{1-\lambda+\lambda\eta\kappa}}_{\nu} a_t. \tag{1.52}
$$

The New Keynesian Phillips curve reads:

$$
\pi_t = \beta\mathrm{E}_t\pi_{t+1} + \psi mc_t, \tag{1.53}
$$

with $\psi \equiv \frac{(1-\omega\beta)(1-\omega)}{\omega}$. Real marginal costs can be expressed as $mc_t = w_t - a_t$. Using (1.52), mc_t can be expressed as:

$$mc_t = \vartheta y_t + (\nu - 1)a_t, \tag{1.54}$$

therefore:

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \psi \left[\vartheta y_t + (\nu - 1)a_t\right]. \tag{1.55}$$

The natural level of output y_t^* is obtained from (1.55) by setting the inflation to zero:

$$y_t^* = \frac{1}{\vartheta}(1 - \nu)a_t, \tag{1.56}$$

the Phillips curve can be rewritten in terms of output gap $x_t \equiv y_t - y_t^*$:

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \psi \vartheta x_t. \tag{1.57}$$

From the IS curve given by (1.51) evaluated at zero inflation one obtains the natural level of the real interest rate:

$$r_t^* = \delta\gamma \left(\frac{1}{\vartheta}(1 - \nu) - \chi\right) \left[\mathrm{E}_t a_{t+1} - a_t\right]. \tag{1.58}$$

The IS curve in terms of output gap reads:

$$\mathrm{E}_t x_{t+1} - x_t = \frac{1}{\delta\gamma}\left(i_t - \mathrm{E}_t \pi_{t+1} - r_t^*\right) \tag{1.59}$$

Model solution

Collect variables into a vector: $z_t = [\pi_t, y_t, i_t, y_t^*]'$. The equilibrium conditions are collected in matrix form:

$$\begin{bmatrix} 1 & -\psi\vartheta & 0 & \psi\vartheta \\ 0 & 1 & \frac{1}{\delta\gamma} & 0 \\ -(1-\phi_i)\phi_\pi & -(1-\phi_i)\phi_x & 1 & (1-\phi_i)\phi_x \\ 0 & 0 & 0 & 1 \end{bmatrix} z_t = \begin{bmatrix} \beta & 0 & 0 & 0 \\ \frac{1}{\delta\gamma} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathrm{E}_t z_{t+1} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_i & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} z_{t-1} + \tag{1.60}$$

$$+ \begin{bmatrix} 0 \\ -\chi(\rho_a - 1) \\ 0 \\ \frac{1}{\vartheta}(1-\nu) \end{bmatrix} a_t. \tag{1.61}$$

In a compact form:

$$B_1 z_t = A\mathrm{E}_t z_{t+1} + B_2 z_{t-1} + C a_t, \tag{1.62}$$

where matrices A, B_1, B_2, C are implicitly defined by comparing (1.61) with (1.62). The rational expectations equilibrium follows:

$$z_t = \Gamma z_{t-1} + \Xi a_t. \tag{1.63}$$

The model is solved by the forward method proposed in Cho and Moreno (2011). For convenience, I define $X_t = (z_t', a_t)'$ and rewrite (1.63) in a compact form:

$$X_t = \varpi X_{t-1} + \Psi \varepsilon_t^a, \tag{1.64}$$

where $\varpi = \begin{bmatrix} \Gamma & \Xi\rho_a \\ \mathbf{0}_{1\times 4} & \rho_a \end{bmatrix}$ and $\Psi = \begin{bmatrix} \Xi \\ 1 \end{bmatrix}$.

Affine term structure model

In setting up the affine term structure model, I follow Bekaert, Cho, and Moreno (2010) and keep the risk premia constant. Given that the paper focuses on intertemporal trade-off and expectations, time-varying risk premia are of secondary importance. The log of the nominal pricing kernel implied by the model outlined in Section II, reads:

$$m_{t,t+1} = \log(\beta) - \gamma\Delta c_{s,t+1} - \pi_{t+1}, \tag{1.65}$$

where $\Delta c_{s,t+1} = \log C_{s,t+1} - \log C_{s,t}$ denotes the consumption growth of savers. Assuming log-normality, the pricing equation for short-term bonds implies:

$$\mathrm{E}_t m_{t,t+1} + \frac{1}{2}\mathrm{Var}_t m_{t,t+1} = -i_t. \tag{1.66}$$

Rewrite (1.65) using the relationship between the consumption of savers and the total output given in (1.48):

$$m_{t,t+1} = \log(\beta) - \gamma\delta\Delta y_{t+1} - \gamma\zeta\Delta a_{t+1} - \pi_{t+1}. \tag{1.67}$$

Define:

$$\Lambda' = \begin{bmatrix} 1 & \gamma\delta & 0 & 0 & \gamma\zeta \end{bmatrix} \Psi,$$

and $\mathrm{Var}\ \varepsilon_t^a = D$, then $\mathrm{Var}_t m_{t,t+1} = \Lambda' D \Lambda$. Given the assumptions, $m_{t,t+1}$ can be written as:

$$m_{t,t+1} = -i_t - \frac{1}{2}\Lambda' D \Lambda + \Lambda' \epsilon_{t+1}, \tag{1.68}$$

and the n-period zero coupon yield reads:

$$\begin{aligned}
y_t^{(n)} &= -\frac{A_n}{n} - \frac{B_n}{n} X_t \\
A_n &= A_{n-1} + \frac{1}{2} B'_{n-1} \Psi D \Psi' B_{n-1} - \Lambda' D \Psi' B_{n-1} \\
B_n &= B'_{n-1} \varpi \\
B_1 &= -e'_3.
\end{aligned}$$

To test if the model-implied term structure recovers the short term real rate, I simulate the model at the parameters given in Table VI for 220 quarters and construct the yield curve. Then, I construct the measure of the real rate $\tilde{r}_t = i_t - y_t^{(10)}$ as in the previous sections and compare it to $i_t - \mathrm{E}_t \pi_{t+1}$ where $\mathrm{E}_t \pi_{t+1}$ is obtained from an AR(1) model. Figure 1.11 shows the histogram of correlations between $\tilde{r}_t$ and $i_t - \mathrm{E}_t \pi_{t+1}$. The median correlation is around 0.7 which supports the validity of $\tilde{r}_t$ as a measure of the short term real rate.

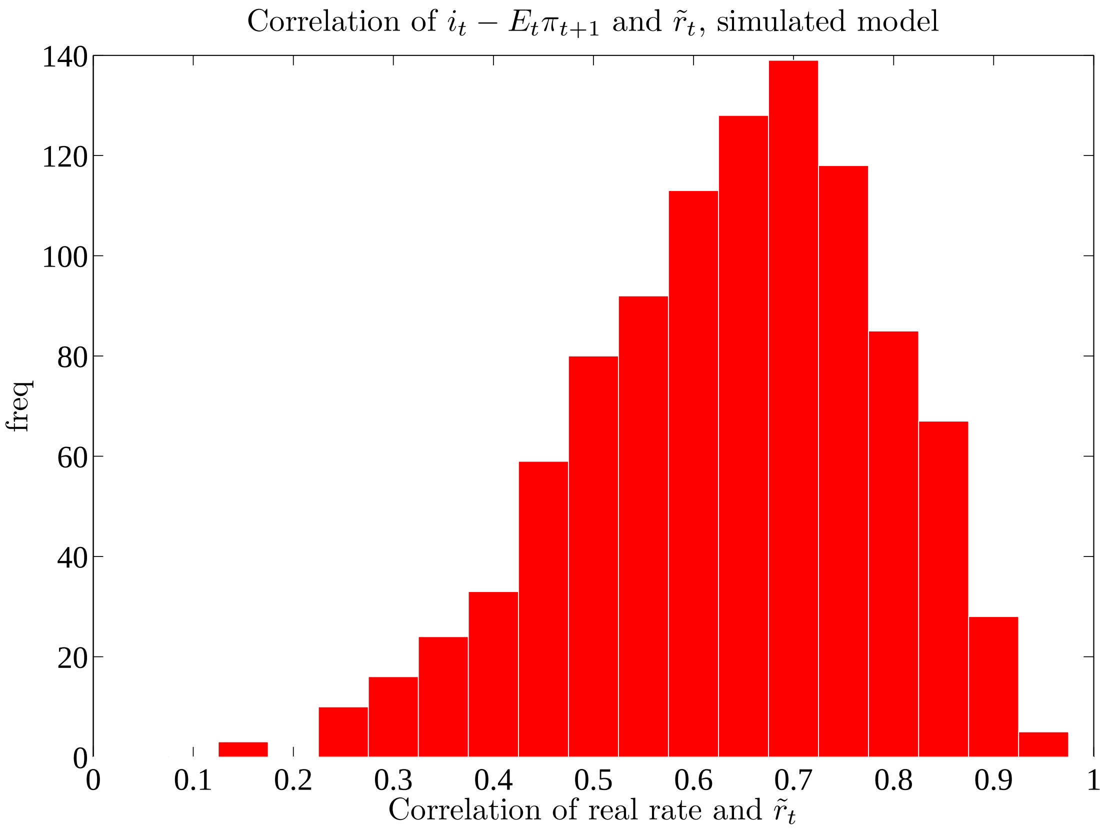


Figure 1.11: Correlation of $\tilde{r}_t$ and real rate in a simulated model.

The figure reports the distribution of the correlation between $\tilde{r}_t$ and the ex-ante real interest rate from a simulated model. $\tilde{r}_t$ is obtained as a difference between the nominal short rate and the ten-year yield. Ex-ante real rate is computed as a difference between short term nominal yield and expected inflation. To obtain the expected inflation, I estimate an AR(1) model on realized inflation and forecast two quarters ahead. The model is simulated for 220 quarters with 1000 replications.

IX.F. Robustness and additional results

Ex-ante real rate: robustness checks

This appendix compares $\tilde{r}_t$ to alternative measures of the real rate variation. First, I compare $\tilde{r}_t$ constructed using the nominal UK yield curve to the real interest rate obtained from the inflation-indexed government bonds. As shown in Figure 1.12, in the period 1985-2011, $\tilde{r}_t$ closely follows the real interest rate from UK inflation-indexed bonds with the correlation $\rho = 0.6$. The shortest continuously available maturity for the real interest rate is 3.5 years. Intuitively, the one-year real rate shall exhibit more variation than the one plotted in Figure 1.12 and therefore it is more suitable for studying the intertemporal trade-off. Second, I compare $\tilde{r}_t$ in the US to the real rate constructed using the inflation expectations from the SPF panel in the period 1970-2011. Figure 1.13 superimposes $\tilde{r}_t$ with the survey-based real interest rate $i_t - \mathrm{E}_t^s\pi_{t+1}$ where i_t represents the

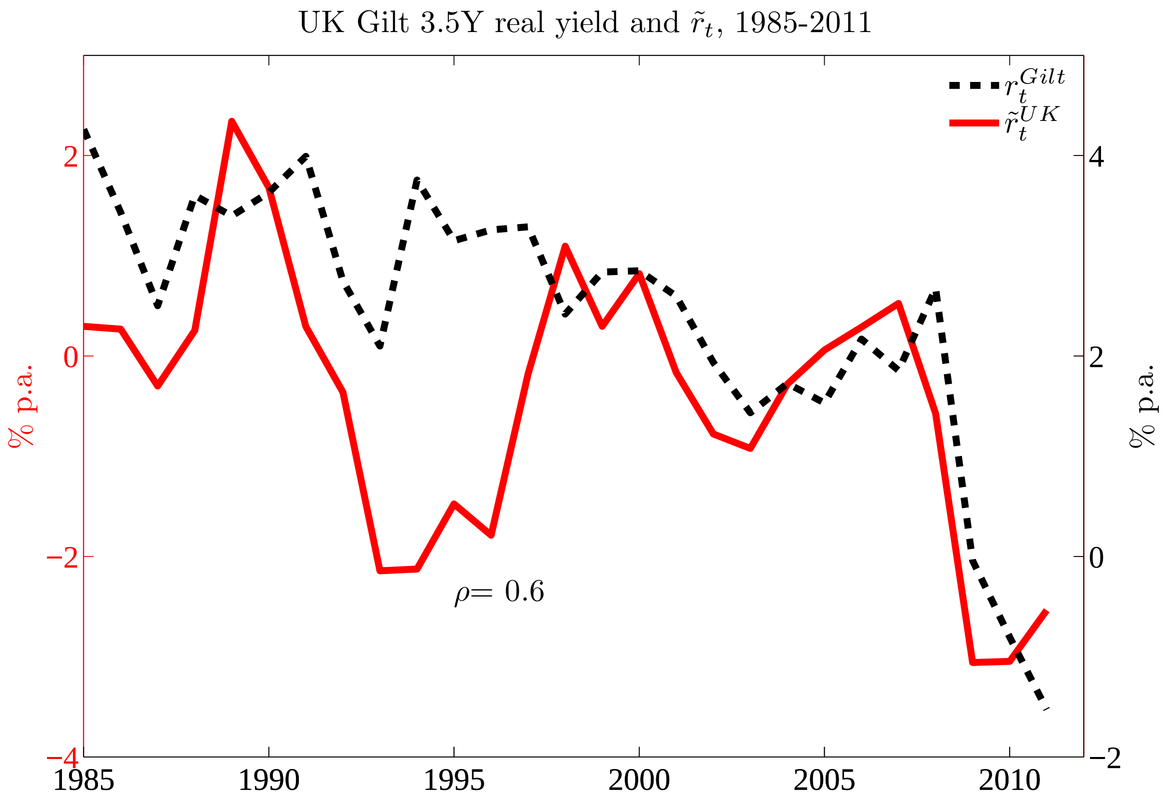


Figure 1.12: Real yield from UK inflation-indexed government bonds and $\tilde{r}_t$, 1985-2011

The figure plots the proxy for the ex-ante real rate $\tilde{r}_t$ which is computed as the difference between the one- and ten-year nominal Gilt. Superimposed is the real yield with maturity 3.5 years denoted by r_t^{Gilt}. Shorter maturities are not continuously available for real yields in the UK. The data are annual. The sample period is 1985 through 2011.

three-month Treasury bill and $\mathrm{E}_t^s \pi_{t+1}$ denotes the survey-based measure of inflation expectations over the next year. Similarly to the real rate obtained from the Livingston survey, both $\tilde{r}_t$ and $i_t - \mathrm{E}_t^s \pi_{t+1}$ closely track each other ($\rho = 0.53$).

Accuracy of the pre-1929 data

The interest rate and consumption data from the period before 1929 are not well-researched. Therefore, there is a potential issue that at least part of the results are driven by the selection of the dataset. To address these concerns, I evaluate the main result, namely the estimate of the EIS using alternative data sources.

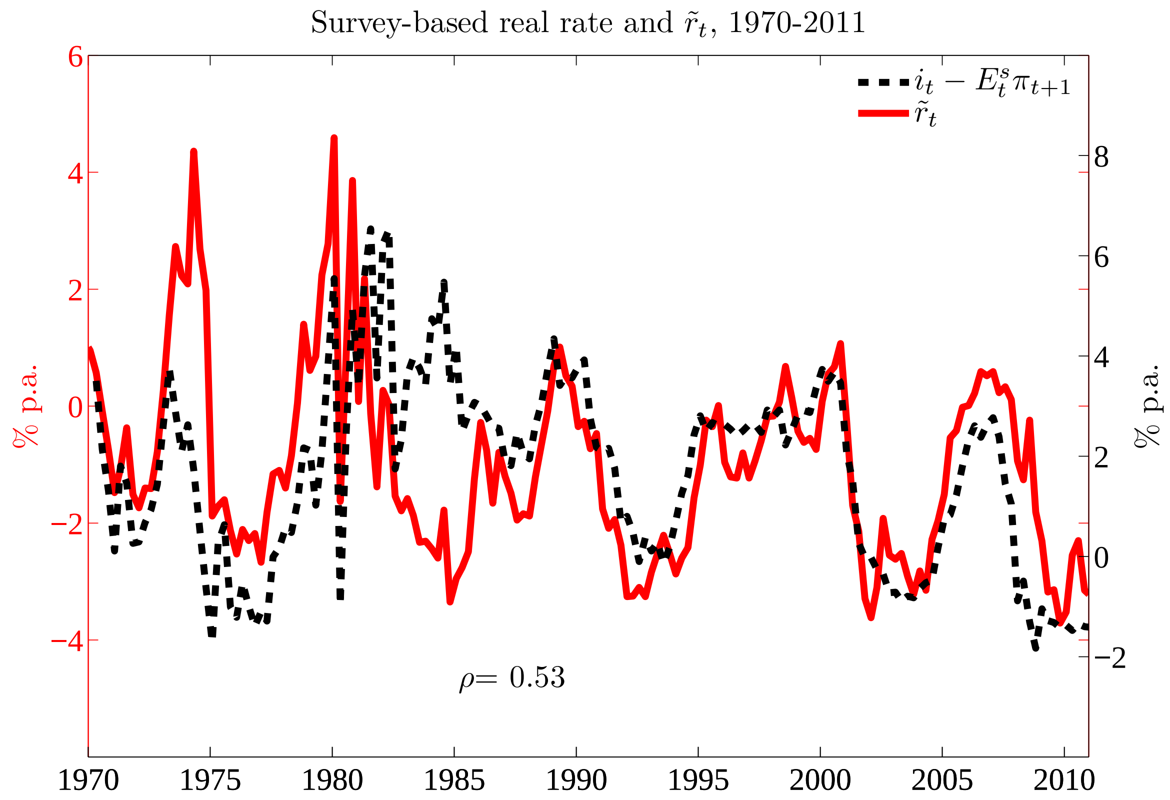


Figure 1.13: Ex-ante real interest rate and $\tilde{r}_t$, 1970-2011

The figure plots the proxy for the ex-ante real rate $\tilde{r}_t$ which is computed as the difference between the three-month Treasury bill yield and ten-year Treasury yield together with the survey-based real interest rate $i_t - \mathrm{E}_t^s \pi_{t+1}$, where i_t represents the three-month Treasury bill and $\mathrm{E}_t^s \pi_{t+1}$ denotes the survey-based measure of inflation expectations over the next year. The survey data are obtained from the SPF panel and the median response is a proxy for the expectations. The data are quarterly. The sample period is Q2:1970 through Q3:2011.

First, I re-estimate the consumption Euler equation using the annual data by Robert Shiller.[37] The sample period is 1888-2009. $\tilde{r}_t$ is defined as the difference between the one-year interest rate and the long term government bond yield, which is a ten-year Treasury note post 1953. The results are quantitatively similar, i.e. the estimate of σ is close to -0.5 for the full sample and it is statistically significant.

Second, between the Civil War and 1920, yields of government bonds are potentially downward-biased due to the fact that government bonds were held as reserves by banks and these could issue bank notes against them. To evaluate the potential bias, I use the non-Treasury yields to construct

[37] http://www.econ.yale.edu/~shiller/data/chapt26.xls

$\tilde{r}_t$ and use it to estimate the Euler equation. The results are quantitatively similar to those using Treasury yields.

Chapter 2

Expecting the Fed

I. Introduction

Separating short rate expectations from risk premia in Treasuries is of importance for policy makers and those seeking to understand the economics of the yield curve.[1] Such decomposition provides insights about how markets perceive the future course of monetary policy, economic activity, inflation and their associated risks. It is also informative about the channels—risk premium versus expectations—through which monetary policy influences the economy.[2] Recent academic research has significantly improved our understanding and measurement of bond risk premia, but still surprisingly little is known as to how investors form expectations about the future path of monetary policy. This focus can be justified with the common assumption of the full-information rational expectations (FIRE) which stipulates that all predictable variation in bond returns comes from risk premia, with expectations formation being of little independent interest. However, analyzing expectations

[1]This chapter is based on the paper with the same title written in collaboration with Anna Cieslak from the Kellogg School of Management, Northwestern University. We thank Snehal Banerjee, Gadi Barlevy, Greg Duffee, Martin Eichenbaum, Douglas Gale, Arvind Krishnamurthy, Scott Joslin, Philippe Müller, Viktor Todorov, as well as Mark Watson and Ken Singleton (discussants) and participants at the SF Fed conference "The Past and Future of Monetary Policy," Financial Economics Workshop at NYU, Arne Ryde Workshop, Chicago Fed, Northwestern Kellogg, Minnesota Mini Asset Pricing Conference, Finance Down Under Conference, Financial Econometrics Conference in Toulouse, and the Red Rock Finance Conference for valuable comments.

[2]See for instance the speech of the former Fed governor Kohn on the importance of this distinction for the policy making (Kohn, 2005).

takes on a new importance as central banks around the world embrace the forward policy guidance.[3]

We start with the observation that lagged information, spanning length of a business cycle, improves predictions of future short rate changes relative to conditioning on the current yield curve alone. This is surprising given that today's cross-section of yields reflects risk-adjusted expectations and therefore, absent additional restrictions, should subsume information relevant for forecasting. We use this observation as a hint to study the properties of private sector's expectations about the future path of monetary policy. Our objective is to assess the degree to which these expectations are consistent with the FIRE or are indicative of informational frictions faced by agents in real time. To directly disentangle the risk premium from short rate expectations, we rely on survey data containing the term structure of private sector's forecasts of the federal funds rate (FFR)—the conventional US monetary policy tool—as well as forecasts of longer maturity yields and inflation. Our results suggest that the view of frictionless rational expectations deviates from the observed behavior of interest rates in several ways.

While survey-based short rate expectations match almost one-to-one the contemporaneous behavior of short-term yields and fed fund futures, these expectations are poor predictors of future short rates except at very short horizons (e.g. Rudebusch, 2002). Consequently, it is relatively easy to identify lagged information, for instance using past term spreads, that improves upon survey forecasts, making short rate forecast errors predictable ex-post. In particular, we find a significant wedge between the ex-ante real short rate implied by the surveys and one that an econometrician would construct under the FIRE assumption. This wedge becomes large in the beginning of NBER-dated recessions, reaching up to -200 basis points: In recessions agents overestimate the ex-ante real FFR compared to the FIRE benchmark.

We construct a measure of expectations frictions as a difference between the survey- and the FIRE-based real fed funds rate. To the extent that expectations of bond market participants are well represented using surveys, this variable reflects the idea that there is information in the time series of monetary policy actions that is not fully impounded in the cross section

[3] The recent speech of the Fed Chairman Bernanke emphasizes the role of forward policy guidance as means to influence the public's expectations about the future path of policy rates (Bernanke, 2011). Recent papers that stress the role of short rate expectations in shaping the reactions to various monetary policy measures during the financial crisis are Bauer and Rudebusch (2011) and Swanson and Williams (2012).

of yields in real time. We label this variable as $MP_t^{\perp}$. One interpretation is that the Fed is able to deliver persistent surprises to the market. Accordingly, we show that $MP_t^{\perp}$ picks up the low-frequency movement in monetary policy surprises identified from high-frequency data (Kuttner, 2001).

The real rate wedge predicts bond excess returns separately from measures of the risk premium in the yield curve, such as the Cochrane and Piazzesi (2005) factor. Its effect is strongest at short maturities (two-year bond), most influenced by the monetary policy, and subsides for longer-term bonds. In a related way, we find that two factors span the predictable variation in realized bond returns across maturities.

With the help of survey data on longer-maturity yields, we obtain a model-free decomposition of annual excess bond returns into a risk premium and an ex-ante unexpected return component. The unexpected return on a two-year bond moves in lockstep with the (negative of) FFR forecast errors with correlation in excess of 0.9. Almost 40% of its variation can be predicted ex-post by $MP_t^{\perp}$. Interestingly, we find that several conditioning variables used to forecast bond returns, especially variables related to the real activity, predict their *un*expected component, comove with $MP_t^{\perp}$, but are essentially uncorrelated with the survey-implied risk premia.

Our evidence suggests that those features of short rate expectations pertain to an environment with an active central bank that itself might adapt its policy rule over time, and are less likely to characterize the data pre-Fed. In the last three decades, we find that agents' forecast errors about the short rate comove closely (with a negative sign) with errors they make when forecasting unemployment, and much less so—inflation.

One concern about the validity of these results is that survey expectations of the federal funds rate may not reflect the true market perceptions of the future evolution of the policy rate. It can be that surveys are noisy, and that forecasters simply anchor their predictions to the current market rates reporting risk-adjusted rather than physical expectations. Thus, in case of such circularity between surveys and market yields, what we identify as expectations frictions could arise from a pure risk premium variation. Indeed, we find a close overlap between survey expectations and expectations extracted from the fed fund futures, which represent a market-wide consensus. We also fail to reject the hypothesis that survey expectations are consistent with those shaping the short end of the Treasury curve. This fact speaks against the hypothesis that noise prevents inference using surveys but it does not address the second concern about circularity.

It is unlikely, though, that forecasters report risk-adjusted predictions for several reasons. First, this argument would imply that the risk premium makes survey forecasts less accurate, and forecast errors more predictable, than they otherwise would be. Using various statistical models with different levels of sophistication, from a simple random walk through a time-varying parameters Bayesian VAR, we find that none is able to outperform surveys in generating more precise real-time forecasts. Moreover, for risk premium to account for our results, one would need to accept that investors charge a highly volatile and implausibly large risk premium (on the scale of several hundred basis points) when investing in short-term and safe rate instruments. While the bulk of our results relies on professional forecasts of the FFR from the Blue Chip Financial Forecasts survey, we uncover analogous results in the Survey of Professional Forecasters comprising different panelists and the T-bill rate predictions. We also report similar properties of expectation errors in the so-called Greenbook forecasts of the FFR, i.e. forecasts prepared by the staff of the Federal Reserve before FOMC meetings. As a last step, we draw on the evidence from money market funds to find that flows in and out of these funds support the expectations frictions interpretation. After controlling for the flight to safety and liquidity episodes, our measure of expectations frictions explains 55% of institutional money market flows and up to 25% of retail flows during a year. Specifically, the inflows gradually increase after monetary policy has been tight and decline after it has been easy, suggesting that money market investors extrapolate from the recent past. The relationship is symmetric in easing and tightening episodes.

Related literature

By studying the role of monetary policy expectations for the yield curve we combine the insights from the term structure literature with the recent developments in macroeconomics that emphasize the role of information imperfections and deviations from perfect rational expectations (see Mankiw and Reis (2011) and Woodford (2012) for overview). The question of how (through which friction) models of monetary policy can generate its lasting effect on the real economy is still debated. One promising route and a growing area of macro research has focussed on information rigidities. Coibion and Gorodnichenko (2011a, 2012) provide evidence that information rigidities present in inflation expectations are consistent with models that relax the FIRE assumption. On the theoretical front, several authors stress the relevance of imperfect knowledge in modeling monetary policy (e.g. Orphanides and Williams, 2005; Woodford, 2010; Angeletos and La'O, 2012). The evidence we collect about the short rate dynamics and expectations is reminiscent of natural expectations introduced by Fuster, Laibson, and Mendel (2010): While many macroeconomic variables have complex hump-shaped dynamics, agents forecast the future using simple models, and thus partially overlook the degree of mean reversion in fundamentals. Building on this literature, our objective is to provide an empirical assessment of expectations frictions faced by bond investors, their link to macroeconomic sources, and relevance for describing the dynamics of yields. Our results have potential implications for the measurement of risk premia and expectations in the curve. Furthermore, they may contribute to the discussion of the channels through which monetary policy impacts the real economy.

We build on the literatures on measuring bond risk premia and on extracting market-based expectations of monetary policy from asset prices. For one, motivated by a widely reported failure of the expectations hypothesis of the term structure, a large body of work has focussed on exploring the risk premium as the source of this violation (Campbell and Shiller, 1991; Fama and Bliss, 1987; Cochrane and Piazzesi, 2005). Consistent with FIRE, a common approach to measuring the risk premium variation is through predictive regressions, i.e. a projection of realized bond returns on a variety of conditioning variables, including the yield curve slope, a set of forward rates and macro variables. Perhaps the most vexing conclusion of the research into bond risk premia is that future bond returns are predictable by variables that have a weak contemporaneous relation with the cross section of yields giving rise to the so-called hidden or unspanned term premia factors. This evidence goes back to Ludvigson and Ng (2009), Cooper and Priestley (2009), and has been formalized in Joslin, Priebsch, and Singleton (2010), Duffee (2011), Barillas and Nimark (2012), and most recently in Joslin, Le, and Singleton (2013).

A parallel literature studies the properties of monetary policy expectations extracted from asset prices (e.g. Rudebusch, 1998; Kuttner, 2001; Cochrane and Piazzesi, 2002; Ferrero and Nobili, 2009). Sack (2004) argues for a time varying but overall small risk premium in the fed fund and eurodollar futures. On the other hand, Piazzesi and Swanson (2008) show that realized excess returns on the fed funds futures are strongly predictable with real variables, implying large countercyclical risk premia on these assets. Our results suggest a close link between hidden factors in term premia and the predictable variation in returns on short-term interest rate instruments that is induced by the way short rate expectations are formed. In particular, the proxy for expectations frictions that we construct can be interpreted as an unspanned monetary policy factor.

A related strand of research uses survey data to study expectations formation in financial markets. In the foreign exchange market, Frankel and Froot (1987) explain the forward premium puzzle with expectations errors, and find that these errors are predictable with past information. Bacchetta, Mertens, and van Wincoop (2009) extend this evidence to other asset classes including stocks and bonds. Using survey data on bond yields in the 1969–1985 period, Froot (1989) shows that predictable forecast errors contribute to the violations of the expectations hypothesis for long-maturity bonds. Piazzesi and Schneider (2011) reach a similar conclusion with more recent data reporting that forecast errors on one- through 30-year Treasury yields are predictable both with the current term spread and a linear combination of forward rates. They argue that risk premia implied by the surveys are more persistent than those obtained with statistical approaches. Building on this literature, we investigate the source of frictions by relating them to the real short rate and to monetary policy. In particular, we document that ex-post there is a wedge between the time series dynamics of the real short rate, and the cross section of yields which captures perceptions of agents about the future short rate path. This real rate wedge is responsible for the predictable variation in the ex-post forecast errors about the policy rate.

Deviations from the FIRE have recently gained prominence in studies of other major asset markets. Singleton (2012) emphasizes the distinctive role of informational frictions and imperfect information in the commodities market to explain the pricing of oil. Using micro-survey data on expectations about inflation, stock returns and house prices, Nagel (2012) relates biases in expectations such as overextrapolation of the recent past to the life-time macroeconomic experiences of individuals. Similarly, in a contemporaneous study, Greenwood and Shleifer (2013) draw on responses from equity investor surveys and flows to confirm the presence of extrapolation in the way investors form expectations about future stock returns. They highlight the discrepancy between the statistical and survey-based risk premium measures.

II. Background

Substantive empirical evidence suggests that variables other than current bond yields have predictive power for future bond returns and, relatedly, future yields. Such a finding has been surprising given that yields today reflect market's conditional expectations of short rates and excess returns to be realized in subsequent periods, and therefore, the current yield curve should contain all information useful for forecasting.[4] This section discusses how expectations frictions can be useful in reconciling the empirical predictability results with this benchmark logic.

Let us consider a realized one-period excess return on a two-period zero coupon bond:

$$rx_{t+1}^{(2)} = -i_{t+1} + 2y_t^{(2)} - i_t, \tag{2.1}$$

where $y_t^{(2)}$ denotes a continuously compounded two-period yield, and i_t is a one-period (short) rate. Rearranging (2.1), the two-period yield can be expressed as:

$$y_t^{(2)} = \frac{1}{2}\left(i_t + i_{t+1}\right) + \frac{1}{2} rx_{t+1}^{(2)}. \tag{2.2}$$

Equation (2.2) is a tautology that follows from the definition of bond returns. Since it holds ex-post, realization-by-realization, it also holds ex ante:

$$y_t^{(2)} = \frac{1}{2} F_t\left(i_t + i_{t+1}\right) + \frac{1}{2} F_t\left(rx_{t+1}^{(2)}\right), \tag{2.3}$$

where $F_t(\cdot) = F(\cdot|I_t)$ is an expectations operator, conditional on all information available at time t, I_t. Importantly, (2.3) holds for any model of expectations formation and for any conditioning information set (e.g. Fama and Bliss, 1987; Fama, 1990).

Most term structure models and tests of the expectations hypothesis assume that $F_t(\cdot)$ is formed under FIRE. Under FIRE, the realized future short rate equals $i_{t+1} = F_t(i_{t+1}) + v_{t+1}$, where the forecast error v_{t+1} is unpredictable by information available at time t. Since the contemporaneous yield curve reflects such expectations, it also summarizes all information relevant for forecasting future interest rates. Thus, under FIRE, a variable can forecast future returns without visibly affecting today's yields only when it impacts expectations of the short

[4] Duffee (2012b) gives a recent comprehensive survey of this literature.

rate and the risk premium in an exactly offsetting manner. Such a cancelation argument has been used to justify why variables that are weakly related to the contemporaneous yield curve can predict future bond returns beyond information that is contained in yields themselves (Duffee, 2011).

An alternative interpretation of this empirical fact, one whose relevance we explore in this paper, builds on the idea that the FIRE may not hold exactly in the data. We note that the identities (2.2) and (2.3) jointly imply:

$$i_{t+1} - F_t(i_{t+1}) = -\left[rx_{t+1}^{(2)} - F_t(rx_{t+1}^{(2)})\right], \tag{2.4}$$

where the left-hand side measures agents' forecast error about the short rate, and the right-hand side—the unexpected return. Through equation (2.3), any forecast error that agents make when predicting the short rate must cancel with unexpected returns that they earn ex-post. Since the cancelation in (2.4) is exact, a variable that predicts forecast errors will by construction have a zero net effect on the current yield curve. This argument holds equally for an n-period bond, for which:

$$\sum_{j=0}^{n-2}\left[i_{t+1+j} - F_t(i_{t+1+j})\right] = -\sum_{j=0}^{n-2}\left[rx_{t+1+j}^{(n-j)} - F_t(rx_{t+1+j}^{(n-j)})\right]. \tag{2.5}$$

It is possible that both effects, i.e. the cancelation of factors within the yield curve and ex-post predictable forecast errors, coexist in the data. We draw on evidence from survey forecasts of the short rate by the private sector and the Fed, from long samples of data to show that deviations from the FIRE have an empirical merit and could account for the observed predictability patterns. This does not necessarily mean, however, that people make obvious mistakes. A broad class of models implies that forecast errors can be predictable without people's behavior being irrational. Such a predictability arises under realistic scenarios: Agents are likely to act under imperfect or noisy information (e.g. Woodford, 2003a). They also are likely not to know the exact monetary policy reaction function but rationally learn about its parameters (Friedman, 1979), which themselves can evolve over time.[5] Alternatively, faced with complex underlying dynamics, they may base their forecasts

[5]Indeed, a growing literature documents that the parameters of the monetary policy rule vary over time, e.g. Primiceri (2005), Boivin (2006), Ang, Boivin, Dong, and Loo-Kung (2010), Coibion and Gorodnichenko (2011b).

on simpler intuitive models that deviate from the truth in a significant way but still imply a small utility loss (Cochrane, 1989).

On a methodological level, the wedge between ex-ante and ex-post introduced by frictions can be interpreted as a wedge between the time series and the cross section of yields within the term structure framework of Joslin, Priebsch, and Singleton (2010). Their framework assumes the existence of macro factors that have predictive properties for future yields in the time series but are unspanned by their current cross section. Importantly, since a subset of state variables does not enter the bond pricing equation in the first place, the model does not require an explicit cancelation between risk premia and expectations to occur within the yield curve.

III. Short-rate expectations

III.A. Measuring short rate expectations with surveys

We use private sector forecasts of the federal funds rate, the main operating target of the Fed, from the Blue Chip Financial Forecasts (BCFF) survey. The survey contains monthly forecasts of the FFR provided by approximately 45 leading financial institutions. The sample of FFR forecasts extends from the inception of the survey in March 1983 through December 2010, spanning a relatively homogenous period for the US monetary policy, during which the FFR was its main operating tool.[6] The forecasts are quarterly averages of the FFR for the current quarter, the next quarter out to four quarters ahead. From the same survey source, we also obtain forecasts of the all-items CPI inflation spanning horizons from the current quarter out to four quarters ahead. Inflation survey is available from June 1984 through December 2010. We use the median forecast across the panelists, because a simple combination of models/forecasters, such as the mean or median, is known to increase the forecast precision (e.g. Stock and Watson, 1998). We confirm this result in our data by studying the persistence in individual forecasters' ability to outperform the median FFR

[6] The forecasts are 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. BCFF does not publish the precise dates as to when the survey was conducted.

forecast. We find that very few forecasters are able to beat the median forecast consistently across different forecast horizon and over longer time spans.[7]

Figure 2.1 plots the time series of survey-based FFR forecasts from two perspectives: Panel a lines up the forecasts for different horizons with the realized FFR at the time when the forecasts are formed; panel b displays the same information in form of conditional term structures of forecasts. Panel a reveals that forecasts closely trace the current realizations of the FFR, suggesting that there is relatively little mean reversion in expectations, i.e. the market expectations of the short rate are formed as if it followed random walk. Panel b indicates that investors systematically underestimate both the degree of monetary tightening and easing.

Which fraction of future changes in the policy rate is anticipated? Focusing on annual horizon, we estimate:

$$\Delta FFR_{t,t+1} = \underbrace{\gamma_2}_{-0.63\ [-2.34]} + \underbrace{\gamma_3}_{1.04\ [3.36]} [E_t^s(FFR_{t+1}) - FFR_t] + \varepsilon_{t+1}^{FE}, \quad \bar{R}^2 = 0.18, \tag{2.6}$$

where $\Delta FFR_{t,t+1} = FFR_{t+1} - FFR_t$ and $E_t^s(FFR_{t+1})$ denotes the survey-based proxy for the expectations about FFR one-year ahead. In Table I, we report analogous results for the forecast horizon h from one quarter to one year. While we cannot reject the null that $\gamma_3 = 1$, we observe significantly negative γ_2 which is due to the zero-lower bound hit in 2008. Excluding the 2008–2010 period gives an insignificant γ_2 and γ_3 close to one (not reported). Given that we cannot reject $\gamma_3 = 1$, regression (2.6) can be interpreted as a decomposition with the variation in ε_t^{FE} reflecting the forecast error. Importantly, estimates in (2.6) show that more than 80% of annual changes in the policy rate is unexpected by the private sector. However, while this failure is usually attributed to the time-varying risk premium, our results suggest that even if risk premium is corrected for (as it is likely to be the case in (2.6)), private sector expectations are able to forecast only a relatively small fraction of future short-rate movements.

[7]Our data allows us to identify a forecaster (an institution contributing to the survey) and trace them over time. To study the persistence in forecast accuracy, we require a forecaster to contribute at least 36 consecutive months to the survey (the samples differ among forecasters). There are 33 contributors who survive this filter. For each forecaster, we measure the ratio of their RMSE relative to the RMSE of the median forecaster. We find that 21% of forecasters are able to achieve a ratio below 1, but only one of them is below 0.95. The distribution of RMSE ratios is strongly skewed to the right with more than 68% of the panelists achieving a ratio of 1.05 or worse.

A forecast error made by the median forecaster about the future policy rate at horizon h is defined as:

$$FE^{FFR}_{t,t+h} = FFR_{t+h} - E^s_t(FFR_{t+h}). \tag{2.7}$$

Panel c of Figure 2.1 shows that forecast errors have nontrivial dynamics over the monetary policy cycle: they are on average negative during easings and positive during tightenings. The most pronounced errors are negative and occur during and after the NBER recessions meaning that forecasters fail most significantly in predicting the timing and the magnitude of the easing. In tightening episodes, the forecasters fail to predict the strength and the pace of interest rate increases. The average error reaches -1.43% and 0.60% at the one-year horizon in easing and tightening episodes, respectively, with standard deviations of 1.37% and 0.88%. As such, the private sector predicts a smaller magnitude of monetary policy actions relative to those that are subsequently realized (more details are in Table XVII in the Appendix). Forecast errors do not seem to be decreasing over time even though the Fed has substantially increased its transparency throughout our sample.[8] A simple regression (not reported) of absolute forecast errors on a time trend confirms that there has been no decline in the errors over time.

III.B. Expectations and the role of lagged information

Survey-based expectations are useful for assessing the discrepancy between the time series and cross-sectional dynamics of the short rate. In particular, we can analyze the degree to which agents' expectations incorporate all past information.

Let us focus on a one-year change in the federal funds rate, $\Delta FFR_{t,t+1}$. To the extent that the Fed's inflation target is slow moving, one can expect $\Delta FFR_{t,t+1}$ to mainly reflect the dynamics of real variables. For instance, over the 1954–2010 period, changes in FFR have comoved strongly with the annual changes in the rate of unemployment with a correlation of -60%. In the post-Volcker sample this correlation strengthened to nearly -70%. Therefore, a variable that predicts real activity is likely to also contain information about future changes

[8] In our sample, there have been several remarkable operational changes that increased the transparency of the Fed. First, in 1994 the Fed started issuing a statement following each FOMC meeting. Starting in March 2002, votes of the committee members are public. In April 2011, the Fed introduced a press conference following every second FOMC meeting. Sellon (2008) finds that the transparency of the monetary policy decreased the prediction errors at short horizons while the prediction errors at longer horizons (one year and more) have not changed.

in the FFR. The slope of the yield curve offers itself as a candidate predictor, given large literature that documents its forecast power for future real activity several quarters ahead (Estrella and Hardouvelis, 1991; Harvey, 1989; Bernanke and Blinder, 1992). We project a one-year change in the FFR on today's and lagged slope defined as the spread between the long- and short-term yields, $S_t = y_t^{(20)} - y_t^{(1)}$:

$$\Delta FFR_{t,t+1} = \alpha_0 + \underbrace{\alpha_1}_{-0.07\ [-0.74]} S_t + \underbrace{\alpha_2}_{0.79\ [6.13]} S_{t-1} + \varepsilon_{t+1}, \quad \bar{R}^2 = 0.37. \tag{2.8}$$

For parsimony and easy interpretation of coefficients, we include only lagged S_{t-1} from one year ago.[9] The predictability implied by the estimates in (2.8) is almost entirely driven by the lagged slope, and significantly higher than the one attained with the survey forecasts in (2.6). The positive sign of α_2 means that high past slope (steep yield curve) is a signal that FFR will increase in the future, which is consistent with the slow mean-reversion of the short rate at the business cycle frequency.

To verify whether agents perceived the dynamics of the short rate in real time in the same way that an econometrician can observe it ex-post, we test if their expectations subsume information in the lagged slope:

$$\Delta FFR_{t,t+1} = \alpha_3 + \underbrace{\alpha_4}_{0.27\ [1.16]} [E_t^s(FFR_{t+1}) - FFR_t] + \underbrace{\alpha_5}_{0.66\ [5.08]} S_{t-1} + \varepsilon_{t+1}, \quad \bar{R}^2 = 0.37. \tag{2.9}$$

We note that in the presence of S_{t-1} the coefficient on the expected path (the first term in (2.9)) drops to 0.27 from 1.04 reported in equation (2.6). This indicates that the two regressors in (2.9) contain common information. However, while under the FIRE α_5 should not be statistically different from zero, the estimates in (2.9) strongly reject this null. This has the important implication that forecast errors are ex-post predictable:

$$FE_{t,t+1}^{FFR} = \delta_0 + \underbrace{\delta_2}_{0.44\ [3.62]} S_{t-1} + \varepsilon_{t+1}, \quad \bar{R}^2 = 0.15. \tag{2.10}$$

In Table II, we report analogous results for different forecast horizons. Private sector forecasts are quite accurate at short horizons but deteriorate rapidly as the horizon increases. This

[9] While allowing more lags improves model specification in terms of BIC and AIC, the improvement is marginal relative to the specification with just one lag and adding multiple lags does not significantly alter our conclusions.

feature is visible in panel B of Table II, where the economic and statistical significance of S_{t-1} for predicting forecast errors increases with the horizon.

III.C. Measuring expectations frictions

The predictability of forecast errors raises the question about economic variables that contribute to the wedge between what agents see ex-ante and what an econometrician finds ex-post. The federal funds rate varies either because of inflation or because the real FFR is not constant through time. To the extent that inflation expectations follow a highly persistent process with a low volatility of shocks (e.g. Neely and Rapach, 2008; Faust and Wright, 2011), by taking changes in the FFR we net out the effect of inflation. In this section, we document that there is a persistent discrepancy between an ex-ante real FFR measured using survey data and ex-ante real rate constructed under the assumption of FIRE. We introduce a measure of expectations frictions that focuses on this aspect of short rate dynamics.

Following the literature (e.g. Laubach and Williams, 2003; Clark and Kozicki, 2004), we define the real federal funds rate, r_t, as:

$$r_t = FFR_t - \pi_t, \tag{2.11}$$

where π_t is the annual inflation, $\pi_t = \log(P_t/P_{t-1})$ and P_t is the level of CPI. We are interested in (2.11) in an ex-ante form:

$$r_t^e = E_t(FFR_{t+1}) - E_t(\pi_{t+1}). \tag{2.12}$$

The literature has proposed different ways of measuring the ex-ante real rate. The most common approach that relies on the FIRE assumption is to obtain the ex-post real rate as $r_{t+1} = y_t^{(1)} - \pi_{t+1}$, where $y_t^{(1)}$ is the one-period nominal yield, and project it on a set of time-t instruments (Fama, 1975; Mishkin, 1981; Yogo, 2004).[10] By taking the nominal yield as given, this approach focusses on approximating the unobserved expected inflation component and abstracts from how expectations about the real economy are formed. Since we are interested in the latter aspect, we instrument not only for π_{t+1} but also for FFR_{t+1}.

[10] Usually, the set of instruments contains π_t, FFR_t, $y_t^{(1)}$ and S_t, but any time-t variable could be a valid instrument.

First, we project r_{t+1} on the following instruments: the year-on-year CPI inflation π_t, FFR_t, a one-year nominal yield $y_t^{(1)}$, an annual change in the rate of unemployment ΔUNE_t, and the term spread S_t.[11] Note that none of the variables is revised or contains forward-looking information. The fitted value from this projection is the FIRE version of the ex-ante real FFR:

$$\hat{r}_t^{e,FIRE} = E_t\left[r_{t+1} \middle| \, \pi_t, y_t^{(1)}, FFR_t, \Delta UNE_t, S_t, S_{t-1} \right]. \tag{2.13}$$

As a second approach, we construct the ex-ante real FFR directly from surveys of professional forecasters:

$$r_t^{e,surv} = E_t^s(FFR_{t+1}) - E_t^s(\pi_{t+1}), \tag{2.14}$$

where as before $E_t^s(\cdot)$ denotes the survey-based expectation. We define a measure of expectations frictions as a difference between $\hat{r}_t^{e,FIRE}$ and $r_t^{e,surv}$, which we call the real rate wedge and denote as $MP_t^{\perp}$:

$$MP_t^{\perp} = \hat{r}_t^{e,FIRE} - r_t^{e,surv}. \tag{2.15}$$

If investors form their expectations in accordance with FIRE, both measures of the ex-ante real rate should coincide, or differ just by a noise component. The empirical properties of $MP_t^{\perp}$ turn out to differ in significant ways from this benchmark.

To construct (2.14), we use inflation and FFR surveys described in Section III.A. This allows us to obtain $MP_t^{\perp}$ at a monthly frequency for the sample period 1984:06–2010:12. Panel *a* of Figure 2.2 superimposes the two real rate estimates, $\hat{r}_t^{e,FIRE}$ and $r_t^{e,surv}$. Panel *b* plots their difference (2.15).

$MP_t^{\perp}$ has several interesting features. First, the wedge becomes consistently negative during the NBER-dated recessions. During recessions investors expect a higher real rate compared to that implied by the FIRE. The magnitude of the deviation is sizable reaching around -200 basis points. Second, $MP_t^{\perp}$ explains a nontrivial part of the variation in ex-post forecast errors about the FFR:

[11] Our results are robust to the selection of instruments. Important for our results is to include lagged information, either in the form of lagged term spread or changes in unemployment.

$$FE_{t,t+1}^{FFR} = \delta_0 + \underbrace{\delta_1}_{0.96\ [6.03]} MP_t^{\perp} + \varepsilon_{t+1}, \quad \bar{R}^2 = 0.39. \tag{2.16}$$

This supports the idea that the wedge captures information that may be omitted from the time-t information set of the forecasters. Third, and in a similar way, $MP_t^{\perp}$ is weakly related to the time-t yield curve. Table III reports a projection of $MP_t^{\perp}$ on the first five yield principal components (PCs). Neither of the first three PCs, which summarize vast part of the yield curve, is statistically significant. The only significant regressor is PC5, albeit we find this result not to be stable across different data sets of zero coupon yields and across sample periods. Together, five PCs account for 19% of the variance in the real rate wedge.

One may wonder whether the weak link between yields and $MP_t^{\perp}$ is an artifact of survey data which may not reflect true expectations of bond market participants. We find, however, that the average survey expectation of the FFR over the next four quarters comoves very closely with the contemporaneous one-year yield, explaining 99% of $y_t^{(1)}$ in levels and 94% in annual changes. We handle the question of survey reliability in Section V in more detail (see also Appendix X.B).

$MP_t^{\perp}$ can be interpreted as an indication that the monetary policy is able to generate persistent surprises relative to the expectations of the public. Therefore, it is useful to compare its dynamics with standard measures of monetary policy shocks. One widely applied measure based on the fed fund futures has been suggested by Kuttner (2001), and further supported by Piazzesi and Swanson (2008) as more robust to the presence of risk premia compared to other alternatives. Kuttner (2001) obtains monetary policy shocks from one-day changes in the fed fund futures around the FOMC announcements. The data is available on his webpage for the period 1989:06–2008:06, which we extend through 2010:12 using the same methodology. Panel *a* of Figure 2.3 plots the daily series of monetary policy shocks. An interesting observation is that monetary policy shocks appear in clusters. Specifically, initially negative surprises are followed by more negative surprises resulting in persistent dynamics. In panel *b* of Figure 2.3, we superimpose $MP_t^{\perp}$ with the time series of cumulative Kuttner's surprises defined as the moving sum of the daily shocks accumulated over eight consecutive FOMC meetings (an approximate number of meetings per year), thus matching the annual horizon of $MP_t^{\perp}$. The cumulative surprises confirm the persistent nature of monetary policy shocks, and also point to a large degree of their comovement with $MP_t^{\perp}$ with correlation of 57%. The largest discrepancies between the two series occur in the early part of the sample. Indeed, before 1994 the Fed was not explicitly announcing changes to its target, which could complicate the identification of monetary policy shocks in that period (Kuttner, 2003). Overall, however, these results suggest that $MP_t^{\perp}$ is related to the persistent component of monetary policy surprises which is not contained in today's market expectations.

IV. Bond excess returns and ex-post forecast errors

This section studies whether expectations frictions could affect the measurement and interpretation of bond risk premia. While a common approach to measuring premia is through predictive regressions of realized returns on a set of conditioning variables, our previous results suggest that part of variation identified in this way may come from the ex-post predictability of forecast errors. This distinction is important as the two channels are economically different. In standard asset pricing models, risk premia reflect the compensation expected and required by investors for the covariance risk of Treasury returns with their marginal utility. Expectations frictions, in turn, are manifest in the ex-post predictability of ex-ante unexpected returns after the risk premium has been corrected for.

In a first step, we notice that the real rate wedge, $MP_t^{\perp}$, has predictive power for the realized excess bond returns. In a second step, we further decompose the realized return into an expected and unexpected part, and study their properties. To summarize the outcome, up to half of the predictable variation in realized bond returns stems from a component that is ex-ante unexpected. The effect, however, is not uniform across maturities: it is strongest at the short end of the yield curve, and subsides as the maturity increases.

IV.A. Predictive regressions of bond excess returns

We estimate standard predictive regressions of bond excess returns across maturities:

$$rx_{t,t+1}^{(n)} = \delta_0 + \delta_1 RP_t + \delta_2 MP_t^{\perp} + \varepsilon_{t,t+1}^{(n)}, \tag{2.17}$$

where $rx_{t,t+1}^{(n)}$ is the annual holding period excess return on a Treasury bond with n years to maturity, and RP_t is an empirical measure of bond risk premia, i.e. of the *expected* component of returns. We use two alternative measures of RP_t from the earlier literature: the linear combination of forward rates proposed by Cochrane and Piazzesi (2005), CP_t, and the cycles factor $\widehat{cf}_t$ from Cieslak and Povala (2011). The CP_t is a commonly used in-sample benchmark for the time-varying bond risk premium. The $\widehat{cf}_t$ can be constructed in quasi real-time by estimating a small number of parameters, it has a stable out-of-sample properties, and as we document below, it does not predict ex-post forecast errors.[12]

Table IV summarizes the results of return forecasting regressions for bonds with maturities of two, three, five, ten and twenty years.[13] Due to overlapping data, we report t-statistics based on Hodrick's reverse regression (rows "t(H)") as well as the Newey-West t-statistics (rows "t(NW)"). Panel A, B1 and C1 report univariate regressions using $MP_t^{\perp}$, and the two risk premium variables, respectively. The main observation is that $MP_t^{\perp}$ is a significant predictor of realized excess returns (panel A), and that predictive power comes mainly from its component that is orthogonal to the contemporaneous yield curve (last row of panel A). The predictive power of $MP_t^{\perp}$ is the most significant, both economically and statistically, at short maturities. In contrast, for both risk premium proxies (panels B1 and C1), the significance of coefficients increases with the maturity, and the explained fraction of returns at the short end of the yield curve is about half of that at the long end. Panels B2 and C2 report the estimates of equation (2.17). The negative sign of the δ_2 coefficient is consistent with the interpretation that lower $MP_t^{\perp}$ anticipates higher bond returns and lower yields in the future. In the presence of $MP_t^{\perp}$, the significance of either $\widehat{cf}_t$ or CP_t remains nearly unchanged, indicating that the real rate wedge captures a new source of predictability.

One may be concerned about statistical biases that arise with long-horizon returns, overlapping data, and artificially splined zero-coupon yield curves that we use above. Therefore, in Table V we repeat the predictive exercise with monthly excess returns on true bond portfolios from CRSP. The estimates confirm our above conclusions. Specifically, $MP_t^{\perp}$ has a negative and highly significant loading, and it dominates the other two predictors in forecasting returns of portfolios with short maturities.

These results suggest that realized bonds returns move around on two factors which represent largely independent sources of their predictability. Indeed, we notice that one can construct two orthogonal factors ($rx_{t+1}^{L}, rx_{t+1}^{S\perp L}$) that span almost the entire variation of realized returns

[12] Cieslak and Povala (2011) decompose the yield curve into long-horizon inflation expectations and maturity-related interest rate cycles. Then, the term structure of cycles is used to separate the risk premium variation from the business cycle variation in short rate expectations.

[13] We obtain zero coupon yields from the constant maturity Treasury (CMT) rates provided by the Fed Board.

across different maturities. The first factor rx^{L}_{t+1} is simply the return on the long term bond (20-year maturity), while the second factor $rx^{S\perp L}_{t+1}$ represents part of the return on the short-term bond (two-year maturity) that is orthogonal to rx^{L}_{t+1}. Using this two-factor decomposition, we find that $rx^{S\perp L}_{t+1}$ is strongly predictable by our measure of expectations frictions, but is unrelated to RP_t proxies. For rx^{L}_{t+1} the reverse holds true. These regressions are not reported in any table for brevity.

IV.B. Decomposing realized bond returns

To explicitly decompose bond returns into an expected and ex-ante unexpected component, we rely on survey forecasts of interest rates from the BCFF survey available from December 1987 through December 2010. The survey contains private sector's predictions of interest rates at different maturities and for horizons of one through four quarters ahead. The panel of participants is the same as for the FFR survey forecasts.

Interest rate forecasts are useful because they allow us to separate in a model-free way an expected part (risk premium) and an unexpected part (forecast error) of realized returns. We focus on the two-year bond return for two reasons. First, for the two-year bond the BCFF data allow us to construct a direct (without approximations) survey-based expected excess return for a one-year holding period. Second, this maturity captures the segment of the yield curve for which we expect the effect of expectations frictions to be most relevant. Using survey forecasts of the one-year yield one year ahead, we obtain a decomposition of the realized excess return into an expected and unexpected component as:

$$rx^{(2)}_{t,t+1} = \underbrace{\left[f^{(2)}_t - E^s_t(y^{(1)}_{t+1})\right]}_{\substack{\text{risk premium} \\ E^s_t(rx^{(2)}_{t,t+1})}} \underbrace{- \left[y^{(1)}_{t+1} - E^s_t(y^{(1)}_{t+1})\right]}_{\substack{\text{unexpected return} \\ rx^{(2)}_{t,t+1} - E^s_t(rx^{(2)}_{t,t+1})}}. \tag{2.18}$$

In equation (2.18), the unexpected return is equivalent to agents' forecast error about the evolution of the one-year rate at the one-year horizon (with a minus sign), $-\left[y^{(1)}_{t+1} - E^s_t(y^{(1)}_{t+1})\right] = rx^{(2)}_{t,t+1} - E_t(rx^{(2)}_{t,t+1})$. This variable, in turn, is strongly correlated with the FFR forecast errors, $\text{corr}\left(FE^{FFR}_{t,t+1}, \left[rx^{(2)}_{t,t+1} - E_t(rx^{(2)}_{t,t+1})\right]\right) = -0.93$.

In Table VI, panel A, we regress each of the two elements on the RHS of (2.18) on $MP^{\perp}_t$ and other time-t predictors. For comparison, we perform a similar exercise using $FE^{FFR}_{t,t+1}$ as the dependent variable, on a sample starting in 1987. The main conclusion is that $MP_t^{\perp}$ predicts a significant fraction of *un*expected returns and of the FFR forecast errors, but has no explanatory power for the expected return component. These regressions are in column (1) of each subpanel of Table VI. Columns (2)–(4) run regressions allowing separate loadings on $\hat{r}_t^{e,FIRE}$ and $r_t^{e,surv}$. While alone each component contributes little to predicting the unexpected return, jointly both become highly significant. In particular, the free coefficient loadings are very close to the 1,-1 restriction that we impose when constructing $MP_t^{\perp}$ in (2.15). Finally, column (6) reports that while $\widehat{cf}_t$ has a strong correlation with survey-based expected return on the two-year bond, it shows no predictability of the unexpected return and of $FE_{t,t+1}^{FFR}$ supporting its interpretation as a expected return.

It is useful to link these results to the recent literature that has emphasized the role of hidden or unspanned factors in driving bond risk premia, i.e. factors that predict returns but are weakly related to contemporaneous yields. While several macro variables have been shown to have this feature, especially those related to the real activity, the economic underpinnings of such factors are still debated. The real rate wedge lends itself for an interpretation as the unspanned monetary policy factor, but rather than to the usual notion of risk premia, it points to an existence of expectations rigidities. When we orthogonalize $MP_t^{\perp}$ with respect to the information in the yield curve, by projecting it on five contemporaneous PCs, the resulting factor is 90% correlated with the original one. Its explanatory power for the forecast error, $FE_{t,t+1}^{FFR}$ increases marginally (by 3 percentage points), in line with the intuition that forecast error predictability should come from variables that are not spanned by the contemporaneous yield curve.

It is worth establishing a link between $MP_t^{\perp}$ and macro variables that have been documented to forecast returns. Beginning with Cooper and Priestley (2009) and Ludvigson and Ng (2009), many authors find that real activity variables help predict excess bond returns beyond the predictability attained with yields or forward rates. This literature also recognizes that real variables are only weakly spanned by the cross section of yields.[14] How do our findings relate to this result? We address this question in panel B of Table VI, by regressing each of the elements on the RHS of (2.18) on two measures of real activity: Chicago Fed National Activity Index (CFNAI) and the annual change in unemployment (ΔUNE_t), respectively.

[14] The common approach to show the lack of spanning is to project a macro variable on yields with different maturities. For real activity measures, the R^2 from this regressions is typically low, suggesting that the cross section of yields does not span the information that a given variable contains.

CFNAI is essentially indistinguishable from the real activity factor constructed in Ludvigson and Ng (2009), and is a version of the Stock and Watson (1999) common real factor. The main observation is that while neither of the real variables has explanatory power for the risk premium part, both are strongly significant predictors of unexpected returns and monetary policy forecast errors. The estimates of bivariate regressions using real activity proxies jointly with $MP_t^{\perp}$ support a weak relationship of those variables with expected returns, but a strong relationship with the forecast errors and unexpected returns. It is unclear whether real macro variables contain new information relative to $MP_t^{\perp}$. While it is possible that our measure of information frictions is imperfect, we note that the coefficient loadings on the macro variables interact with $MP_t^{\perp}$. In Figure 2.4 we superimpose $MP_t^{\perp}$ with ΔUNE_t and CFNAI_t, showing that there is comovement between the series (correlation of -0.35 and 0.36, respectively), and the troughs in $MP_t^{\perp}$ precede those in real activity.

V. Additional evidence on the quality of surveys

In this section, we summarize evidence from additional data sources that provide different angles of assessing the expectations formation process in the yield curve. First, we ask how easy it is to outperform survey forecasts of the federal funds rate with statistical models in real time. Second, we compare surveys with market-based forecast of the FFR from the fed fund futures. Third, we analyze whether internal FFR forecasts of the staff at the Federal Reserve Board are subject to expectations frictions similar to those of the private sector. Fourth, we link the FFR forecast errors that agents make when forecasting macro variables, i.e. unemployment and inflation. Fifth, we perform statistical tests for the presence of information rigidities in the FFR forecasts consistent with sticky and noisy information models. Finally, we discuss evidence from money market flows.

V.A. Do statistical models outperform surveys in real time?

This section compares forecast accuracy of surveys with several statistical models of the short rate estimated in real time. The main results are in panel A of Table VII. Given ample evidence that simple methods of forecasting interest rates often work best in real time (e.g. Duffee, 2009; Wright, 2011), we report naive forecasts assuming the FFR to follow a random walk (row 2), and two univariate specifications: an AR(2) (row 3) and an AR(p) allowing up to 16 quarterly lags which are selected dynamically with the BIC from all possible lag combinations (row 4). We additionally consider three multivariate specifications (rows 5 through 7): a recursive VAR(2) estimates obtained with OLS (row 5), and two Bayesian VARs: a constant parameters VAR(2) with a Minnesota prior (row 6) and a time-varying homoscedastic VAR(2) with time varying parameters in the spirit of Primiceri (2005) (row 7). All VARs are second order and include three variables: CPI inflation, unemployment and the FFR. All models are estimated recursively on an expanding window with a burn-in period of 73 quarters, with the out-of-sample forecasts constructed for the period of 1983:Q1 through 2010:Q4. The out-of-sample period corresponds to the period for which the relevant survey data are available.

Across all forecast horizons, surveys provide the lowest RMSE by a wide margin (row 1), followed by the autoregressive model with a fixed number of lags (AR(2)), and by the random walk. For instance, compared to the forecast error from the AR(2), the relative error made by survey forecasters ranges from 63% at a one-quarter horizon to 92% at four quarters. Importantly, also more sophisticated methods, including time-varying Bayesian VARs, fail to match the precision of the FFR survey forecasts in real time. These results for the FFR resonate well with the finding that, at least in the recent data, surveys tend to outperform statistical forecasting methods, as documented for inflation forecasts by Ang, Bekaert, and Wei (2007b) using post-1985 out-of-sample period.[15]

V.B. Market-based forecasts

In panel B of Table VII, we compare forecast errors made by the median survey panelist to the ones implied by the fed fund futures. Historical futures data are available from Bloomberg starting from 1988:12 for contract horizons up to six months. We match end-of-month futures data with the monthly survey forecasts.[16] Clearly, futures-based forecasts of the FFR differ

---

[15] While surveys forecasts are hard to beat with statistical models at the short maturity range, it is possible to outperform them in real time at longer maturities (two years and above). This is consistent with the evidence that statistical models estimated in real time predict excess returns on bonds with long maturities better than surveys. For instance, at a forecast horizon of one-quarter a dynamic affine term structure model with three factor generates lower RMSEs than surveys at maturities of two years and above (the outperformance is marginal for the two-year bond, and increases with the maturity). However, at the forecast horizon of four-quarters ahead, surveys produce uniformly lower forecast errors for all maturities, and their outperformance is most visible at the short end of the yield curve. We thank our discussant Ken Singleton for making this point.

[16] Futures data have been used in earlier studies by Kuttner (2001), Piazzesi and Swanson (2008), Gurkaynak, Sack, and Swanson (2005), among others. The comparison of survey and futures forecasts is

from the physical forecasts by the presence of a risk premium. Using surveys, we obtain an estimate of the futures risk premium that is on average four basis points for the six-month contract with a standard deviation of 16 basis points. Given the small magnitude of the risk premium, the forecast errors implied from the futures (i.e. the negative of the realized futures returns) are highly correlated with these from surveys, with correlation coefficient of 0.89 at a three-month horizon and 0.93 at a six-month horizon. The futures-based RMSEs for the three- and six-month ahead forecasts are marginally lower relative to the surveys, by three and two basis points respectively, but for the six-month horizon we fail to find a statistically significant difference between these two sources of FFR predictions.

One interpretation is that the median survey response represents quite well market-wide expectations of the short rate. Another interpretation is that survey respondents simply anchor their forecasts at the current market rates, and thus report risk-adjusted rather than physical expectations.[17] This latter hypothesis is unlikely to hold true for several reasons. First, the evidence above tells us that in real time it is hard to beat survey forecasts with statistical models of the physical short rate dynamics. Therefore, the risk premium that forecasters potentially include when forming their expectations, if any, should not be a significant confounding factor. Second, we obtain very similar estimates of risk premia in short-term interest rates to those in the fed fund futures when using expectations of different survey respondents (Survey of Professional Forecasters, SPF) and for other interest rates (3-month T-bill). These estimates confirm that risk premia at the short end of the yield curve are small relative to the overall variation in short-term rates, and are volatile around zero. Contrary to the realized returns, they also systematically decline before and at the beginning of recessions, consistent with the role of short maturity instruments in liquidity and safety provision.

---

necessarily imperfect because futures are settled based on the average FFR that prevails during the contract month, while the forecasters predict average quarterly FFR rates. To make the setup comparable, we use monthly data, and calculate the survey forecast error with respect to the monthly average of the FFR that prevails at time 3, 6, 9 and 12 months from the time of the forecast. Survey forecast errors when using either quarterly or monthly FFR averages are very highly correlated, with correlations 0.94, 0.98, 0.99 and 0.99 for one through four quarters ahead, respectively.

[17] Private conversations with some of the prominent survey participants suggest that forecasters understand very well the difference between the physical and risk neutral dynamics and do not anchor their forecast to the latter, but rather use sophisticated models and judgement to form their predictions.

V.C. Expectations by the Federal Reserve's staff

Before each FOMC meeting, the staff of the Federal Reserve prepares their own forecasts of the FFR from the current quarter up to five quarters ahead. The forecasts are published in Greenbook with a five-year lag and available in the financial assumptions files on the Philadelphia's Fed website. The Greenbook has several useful characteristics. First, the staff at the Fed has extensive access to economic data resources when forming their predictions. The publication lag together with the Fed's ability to observe the current expectations of the market participants, and their possibly better understanding of the policy rule can lead to information asymmetries between the private sector and the policy makers (Romer and Romer, 2000). Second, forecasts of the staff are unlikely to be influenced by subjective, worst-case scenario considerations that characterize the forecasts of the FOMC members (Romer and Romer, 2008; Ellison and Sargent, 2010). Finally, because the names of individual members of the staff are not revealed, reputational concerns are a potentially lesser issue compared with the private sector surveys. In sum, Greenbook forecasts could be viewed as an upper bound on the short rate predictability that agents are able to attain.[18]

In Table VIII we provide cross-correlations between four-quarter ahead forecast errors of the Fed staff and the private sector for the FFR. Despite differences in the characteristics of the forecasters' panel and in their access to information, the FFR forecast errors of the staff and the private sector have a correlation reaching 88%.

In Table IX, we revisit the question whether expectations of the Fed staff are closer to the FIRE compared with those of the private sector. Specifically, we use their forecasts of the FFR at the frequency of the FOMC meetings to replicate the regressions reported in Table II in the available Greenbook sample 1983–2006. The forecasts of the staff are surprisingly similar to those of the private sector: While the economic and statistical significance of past information is somewhat lower, the lagged values of the slope S_{t-1} predict forecast errors, and, in the presence of the Greenbook expectations, continue to be significant predictors of short rate changes. For instance, at the four quarter horizon, S_{t-1} explains 20% of the variation in the Greenbook forecast errors, with a coefficient loading of 0.53 (Newey-West

[18] We abstract here from the fact that the number of staff members preparing the forecasts may be smaller than the number of participants in the private sector surveys, and therefore the Greenbook forecasts may not equally benefit from the forecast errors averaging which tend to increase the forecast precision.

t-statistic of 3.62). This evidence indicates that, even with access to extensive information, it is inherently difficult to predict the policy rate in real time.

V.D. Comovement between forecast errors of FFR and macro variables

It is known that a Taylor-type monetary policy rule that includes inflation and unemployment describes the path of the short rate quite well. Therefore, we expect to attribute at least a part of the FFR forecast errors to the forecast errors that agents make about macro variables themselves.

Starting from a basic Taylor rule, according to which monetary policy reacts to inflation and unemployment, we can decompose the monetary policy forecast errors into two components corresponding to macro variables. To this end, we regress $FE^{FFR}_{t,t+h}$ at a given horizon on the corresponding inflation and unemployment forecast errors, $FE^{CPI}_{t,t+h}$ and $FE^{UNE}_{t,t+h}$:

$$FE^{FFR}_{t,t+h} = \gamma_0 + \gamma_1 FE^{UNE}_{t,t+h} + \gamma_2 FE^{CPI}_{t,t+h} + \varepsilon_{t,t+h}. \tag{2.19}$$

Private sector expectations of macro variables are obtained from the quarterly Survey of Professional Forecasters (SPF), which provides a term structure of forecasts at horizons corresponding to those for the FFR. Table VIII reports the cross correlations of the variables involved in (2.19) for the Fed staff and private forecasts. The important observation is that while the FFR errors are relatively weakly correlated with errors on inflation, they comove strongly (and negatively) with those on unemployment.

To establish a more formal link, equation (2.19) cannot be estimated with OLS because forecast errors on macro variables are likely to be correlated with the innovations to errors on the FFR. Therefore, we estimate (2.19) with instrumental variables, using contemporaneous oil shock and lagged values of the CFNAI as instruments.[19]

Table X summarizes the IV regressions for private sector forecasts at horizons three and four quarters ahead. We consider two sample periods: the full sample (1983–2010) and the sample ending in 2006 to make sure that the results do not depend on the spike in the

---

[19]Following Coibion and Gorodnichenko (2011a), we define the oil shock as the residual from an AR(2) model estimated on the four quarter changes in the oil price. The data is obtained from the FRED database. This variable is a valid instrument since it is uncorrelated with the lagged information and orthogonal to shocks to the monetary policy forecast errors.

unemployment during the recent crisis. The key finding is that monetary policy forecast errors comove strongly with those on unemployment and much less with those on the CPI inflation. $FE^{CPI}_{t,t+h}$ is only marginally significant in the pre-crisis sample and is a small contributor to the overall variation in $FE^{FFR}_{t,t+h}$, while $FE^{UNE}_{t,t+h}$ is significant at the 1% level across both samples and forecast horizons. The $\bar{R}^2$ from the regressions indicate that about 40% of the sample variation in the monetary policy forecast errors can be related to macro sources. The results also suggest a causal relationship: Private sector expectation errors about monetary policy arise, at least partially, from their errors in forecasting the path of unemployment, and real activity more generally.[20]

V.E. Testing information frictions

In this section, we test whether the predictability of ex-post FFR forecast errors we document above can be justified within stylized rational models with information rigidities. Specifically, expectations can deviate from the FIRE if rational agents face informational frictions such as noisy information as in Woodford (2003a) or information stickiness as in Mankiw and Reis (2002). Coibion and Gorodnichenko (2011a) show that in these models, the average (across agents) ex-post forecast error should be predictable with a positive sign by the average forecast revision at the corresponding horizon. Accordingly, the baseline test can be performed by estimating:

$$FE^{FFR}_{t,t+h} = \beta_0 + \beta_1 \left[E^s_t(FFR_{t+h}) - E^s_{t-1/4}(FFR_{t+h})\right] + \varepsilon_{t+h}, \tag{2.20}$$

where under information frictions $\beta_1 > 0$. The results of estimating (2.20) are reported in column (1) of Table XI for horizons h from one through three quarters.[21] The coefficient β_1 is positive and statistically significant across h, supporting the hypothesis that forecasters act under information frictions. Forecast updates alone explain up to 17% of the variation in ex-post forecast errors, and their statistical significance is strongest at short horizons.

[20]Interestingly, in unreported results we fail to establish a similarly strong relationship for the forecasts of the Fed staff. This is intuitive to the extent that the Fed believes monetary policy can influence the path of macro variables.

[21]For the FFR we report estimates of (2.20) using the forecast error and forecast revision of the median forecaster to be consistent with the previous results. The results are essentially identical when using means. The mean and median forecast errors and updates are more than 0.99 correlated with each other at corresponding horizons.

The models of information frictions summarized by (2.20) imply that forecast updates should account for the entire predictable variation in ex-post forecast errors. In columns (2) and (3) of Table XI, we augment regression (2.20) with variables that we have found to contain information about the FFR forecast errors, $MP_t^{\perp}$ and lagged slope S_{t-1}, respectively:

$$FE_{t,t+h}^{FFR} = \beta_0 + \beta_1 \left[E_t^s(FFR_{t+h}) - E_{t-1/4}^s(FFR_{t+h}) \right] + \beta_X X_t + \varepsilon_{t+h}, \tag{2.21}$$

where β_X represents the loading on each of these additional factors. The results of the extended test suggest that lagged yield curve information has explanatory power beyond forecast updates, which is increasing with the forecast horizon. For instance, at the three-quarter horizon, $MP_t^{\perp}$ raises $\bar{R}^2$ of the regression by 29% relative to the baseline case (2.20), and is highly statistically significant (t-statistic of 6.4). Similar results, albeit somewhat weaker, pertain to the lagged term structure slope. The lower significance of the slope is consistent with the fact that it partially reflects risk premium variation which should not predict forecast errors.

One way to assess whether frictions that we document reflect a more general feature of the expectations formation is to study their explanatory power for forecast errors about macro variables other than the FFR. To this end, we estimate equations analogous to (2.21) for the forecasts of unemployment and CPI inflation. In constructing both we use unrevised data and the forecasts from the SPF survey. The evidence in favor of expectations frictions is particulary strong for unemployment, where the amount of predictable variation in ex-post forecast errors reaches $\bar{R}^2$'s up to 29%. We notice that our measures of frictions are statistically and economically significant, contributing a large fraction to the explained variation. Since it is hard to argue that forecast errors about unemployment are confounded by risk premium effects, this evidence provides additional support to the importance of expectations frictions present in the yield curve.

V.F. Evidence from money market flows

In addition to survey forecasts, the relative merit of the expectations frictions interpretation can be further assessed through the lens of investors' actions. Since our results pertain to the way people form expectations about the short rate, we should be able to trace the potential effect of frictions to positions investors take in short-term interest rate instruments, such as money market funds. We obtain monthly values of net assets in the US money market funds from the H.6 statistical release of the Fed Board. We define a proxy of annual money market flows as the year-over-year log change in the funds held in the money market funds, and denote this variable as $\text{Flows}_{t-1,t}$. We distinguish between retail and institutional funds, each representing 24% and 68% of the total money market fund assets, respectively, because these two components feature somewhat different flow patterns that are interesting in our context.[22]

If investors were fully forward looking and able to exploit the predictable variation in the short rate in real time, we would expect their money market allocations to increase in anticipation of a higher short rate (tightening) and decrease otherwise (easing). To the extent that a higher $MP_t^{\perp}$ is a signal that the short rate will increase in the near future, we should observe its positive correlation with the flows. Figure 2.5 suggests that the opposite is in fact true in the data: the flows have a highly negative contemporaneous correlation with $MP_t^{\perp}$. One way to explain the sign is that money market flows simply reflect the flight-to-safety episodes that coincide with monetary policy easings. Contrary to this intuition, the relationship between flows and $MP_t^{\perp}$ is symmetric rather than one-sided, i.e. it holds both in easing and tightening episodes. Another possibility, consistent with the expectations frictions story, is that to a large degree market participants extrapolate from the recent past and thus do not fully accommodate the mean reversion in the short rate.

In panel A of Table XII we report predictive regressions of the changes in FFR and the FFR forecast errors, respectively, from t to $t+1$ on the money market flows in the year ending at time t. The regression coefficients are negative for both retail and institutional flows, with a stronger statistical significance in the latter case. Specifically, increased institutional flows are associated with a subsequent monetary policy that is easier than initially expected (i.e. the FFR forecast errors observed ex-post are low or negative).

In panel B of Table XII, we provide three sets of regressions. In column (1), we establish that flows are high when the monetary policy is unexpectedly easy during the same year, as measured by $MP_t^{\perp}$. For instance, the standardized coefficient loadings of institutional

[22] As of 2010 year-end, the total net assets in money market funds in the US reached about 2.8tr USD, making it one of the largest asset classes, according to the 2012 Investment Company Factbook. For these two classes, respectively, the percentage flows have a mean of 5% and 14%, standard deviation of 13% and 16%, a minimum of -32% and -30%, and a maximum of 29% and 49%. While part of these fluctuations are simply due to the changes in money market rates, it is likely to be a small fraction of that due to net flows, given the extent of the overall variation in the net asset values.

flows on $MP_t^{\perp}$ is -0.75 (t-stat $= -8.1$, $R^2 = 0.56$), meaning that a one standard deviation decrease in $MP_t^{\perp}$ coincides with an increase in institutional money market assets of nearly 12%. A similar relationship, yet slightly weaker, holds true for the retail funds. In column (2), we augment $MP_t^{\perp}$ with a number of contemporaneous controls for safety and liquidity.[23] In the presence of controls, the economic and statistical relationship between the flows and $MP_t^{\perp}$ remains largely unchanged as visible in the reported coefficients. In column (3), we predict flows over the subsequent year (from t to $t+1$) with the term structure slope S_t. The sign of the loading on S_t in column (3) is negative meaning that the flows into money markets are high (low) if the slope has been flat (steep) in the recent past. We note that a high slope is associated with an easy monetary policy. One interpretation is that when the FFR has been high in the past agents expect high interest rates to continue to prevail in the future. As such, the direction of money market flows is in line with our evidence about expectations formation in that expectations do not fully reflect the lagged information that is present in the physical dynamics of the short rate.

These results are interesting in the context of the findings of Piazzesi and Swanson (2008). Specifically, Piazzesi and Swanson (2008) show that the realized returns in fed fund futures are predictable with the net open interest of speculators in *eurodollar* futures. This suggests that certain market participants are able to exploit the predictability in the short rate at the expense of others. While we are able to replicate this result in our sample, i.e. net open interest of speculators in eurodollar futures predicts ex-post FFR forecast errors, we fail to establish an analogous link with the net open interest of traders in the fed fund futures market itself. Open interest in the fed funds futures appears unrelated to the predictable variation in the errors. Together with the evidence suggesting the presence of frictions for the cash money market investors, these results may point towards market segmentation, whereby a particular specialized segment of the derivatives market is able to correctly recognize and exploit the expectations frictions. While this topic deserves a deeper investigation, we leave it to our subsequent research.[24]

[23]We include the market-wide liquidity of Pastor and Stambaugh (2003), the value of funding liquidity of Fontaine and Garcia (2011), noise illiquidity measure of Hu, Pan, and Wang (2013), and stock market volatility series VXO. In several other specifications that are unreported in any table we have extended the set of controls with real activity indicators (CFNAI) and have found essentially unchanged results.

[24]We note that while less developed and smaller than the Eurodollar, the fed funds futures market has been expanding rapidly in the last decade reaching 50% of the size of the Eurodollar market in mid-2004 as measured by the nominal open interest. At the same time, after Piazzesi and Swanson (2008) results were

VI. Short rate dynamics pre- and post-Fed

The sample for which survey data are available is relatively short and therefore restrictive. In this section, we explore the predictability of annual short rate changes in the sample 1875–2011 which spans different monetary policy and institutional settings. The goal is to study how predictable short rate changes are by comparing two perspectives: (i) a time-series perspective that uses a highly stylized statistical model of short rate dynamics and (ii) a cross-sectional perspective that extracts predictive information from the cross-section of yields. Absent time-varying risk premia and/or deviations from the FIRE, the time-series and cross-sectional predictability should, in general, coincide.[25] We find, however, that the cross-section appears to contain significantly less information about future rates than an unconstrained time series model. This finding characterizes the data after the founding of the Fed, but not before.

We start with a deliberately simple empirical model in which short rate changes are linear in lagged values of the short rate itself:[26]

$$i_{t+1} - i_t = \alpha_c + \sum_{j\in 0,1,2,4} \alpha_j i_{t-j} + \varepsilon^i_{t+1}, \tag{2.22}$$

where i_t denotes the short rate and time is measures in years. A formal lag selection procedure (reported in Table XV the Appendix) chooses a low number of lags up to four years back. Accordingly, for simplicity, equation (2.22) includes four lags up to four years.[27]

We are interested in both the degree of predictability and the stability of this model. To cover different monetary policy regimes, we consider quarterly data from 1875 through 2011, divided into five subperiods marked by important institutional changes: (i) pre-Fed (1875–

made public in 2004, we have not observed a weakening of the link between the realized fed fund futures returns and the open interest in the eurodollar market.

[25]As we discuss above, this statement does not hold true if there is an exact offsetting between risk premia and expectations within the yield curve—a special case described by Duffee (2011, 2012b).

[26]Linear models of the term structure of this type have been studied in the earlier literature (Modigliani and Sutch, 1966, 1967; Sargent, 1972). Modigliani and Shiller (1973), Mishkin (1980) discuss the conditions under which this type of models is consistent with rational expectations.

[27]By selecting a low number of fixed lags, we avoid issues with collinearity. Table XV in the Appendix re-estimates (2.22) with optimally selected lags using BIC leading to very similar conclusions. Alternatively, we have also implemented a polynomial distributed lag model and found that the it places a statistically and economically significant weight on distant lags.

1913), (*ii*) post-Fed – pre-Accord (1914–1951), (*iii*) post-Accord – pre-Volcker (1951–1979), (*iv*) post-Volcker (1984–2011), and (*v*) post-Accord (1951–2011). Our data is obtained from the Global Financial Database and comprises the long-term government yield, y_t^{LT}, and the three-month rate which we use as a proxy for the short rate, i_t.[28]

Panel A of Table XIII summarizes the results for equation (2.22). There are two main observations. First, one-year changes in i_t are highly predictable, except for the special period of interest rate controls and wars (1914–1951). Second, higher-order lags of the short rate carry significant information about future rates relative to a restricted model that excludes lags, increasing the $\bar{R}^2$ by an order of magnitude. This pattern pertains to the period post-Fed Accord, but is largely absent before. For instance, our simple model explains up to 40% of short rate changes in the pre- and post-Volcker samples. The importance of specific lags varies over time.

Switching to the cross-sectional dimension, current yields contain a footprint of real-time market expectations of short rates going forward. In particular, the term structure slope arises as a linear combination of expected short rate changes and time-varying risk premia. The presence of risk premia confounds the direct measurement of short rate expectations as formed by the market. If, however, the model in (2.22) is a good representation of the physical short rate dynamics, we can use the following projection to decompose the slope into the expected short rate change and the risk premium:

$$S_t = \delta_c + \sum_{j \in 0,1,2,4} \delta_j i_{t-j} + RP_t^S, \tag{2.23}$$

where the slope is defined as $S_t = y_t^{LT} - i_t$ and RP_t^S is the regression residual.[29]

[28] The short rate is constructed as the three-month commercial paper before 1934 and the three-month T-bill rate afterwards. T-bills become available in late 1920s and thus do not cover the period before the Fed. However, using T-bill or commercial paper rate leads to very similar results for the overlapping period, and therefore should not bias the results. In the pre-Fed period commercial paper was considered to have government credit risk. Following the literature, we use quarterly data by smoothing monthly series with a three-quarter moving average because there is evidence that the short rate was highly seasonal prior to the founding of the Fed in 1914 (e.g. Miron, 1986).

[29] Note that due to the data availability, our measure of the slope corresponding to one-year change in the short rate is imperfect. Optimally, one should use the difference between one-year forward rate and the short rate, which are not observed. In using the slope between the long-term and the three-month yields we make the implicit assumption that expectations of short rate changes move on a single mean-reverting factor.

Panel B of Table XIII reports the estimates of (2.23) in different subsamples. The results reveal that a linear combination of past short rates explains a vast fraction of the slope variation, between 68% and 96%.

Assuming that projection (2.23) is the correct model for extracting short rate expectations embedded in the yield curve, we should observe that: (i) the fitted value $\hat{S}_t$ from the regression predicts future short rate changes and (ii) the fitted residual $\widehat{RP}_t^S$ predicts future excess bond returns but not short rate changes, i.e. all information about future short rates that the slope contains is subsumed in $\hat{S}_t$. To verify the latter, panel C of Table XIII summarizes predictive regressions of one-year excess return on a ten-year bond[30] using $\hat{S}_t$ and $\widehat{RP}_t^S$. Across subsamples, the only significant predictor of excess returns is $\widehat{RP}_t^S$ which suggests that lags of the short rate distinctly absorb the part of slope variation, $\hat{S}_t$, related to short rate expectations. Note that in the pre-Fed period marked by the Gold standard, bond excess returns are not predictable. For comparison, panel C also reports analogous $\bar{R}^2$'s (in brackets) obtained with the Cochrane-Piazzesi (CP) factor in the available 1952-2011 period for which the factor can be constructed. The return predictability by $\widehat{RP}_t^S$ is on par with that of the CP factor.

The key insight from decomposing the slope is that, in the period characterized by an active Fed, the amount of predictable variation in the short rate implied by the cross section is in the range of 2% to 17% (panel D) and is significantly lower than the predictability that can be achieved with the time series model (panel A) of up to 40%.

One criticism of the approach above is that by just using the slope we under-represent the cross-sectional information in yields. Do distant lags of the short rate continue to matter if we condition on the complete information in today's yield curve? To answer this question, we extend the predictive regression (2.22) by including up to six principal components (PCs) of yields. We consider the longest possible period 1952-2011 for which we can obtain a cross-section of five zero coupon yields in the Fama-Bliss data. We augment these data with the three-month T-bill rate to capture the variation at the short end of the curve:

$$i_{t+1} - i_t = \beta_0 + \sum_{j=1}^{N} \beta_j PC_t^j + \delta_1 i_{t-\frac{lag1}{4}} + \delta_2 i_{t-\frac{lag2}{4}} + \varepsilon_{t+1}^{PC}, \tag{2.24}$$

[30]Excess return on a ten-year bond is constructed from total return indices on the ten-year bond and the three-month Tbill, both obtained from the GFD database.

with $lag1$ and $lag2$ selected optimally using the BIC criterion considering all lag combinations up to 16 quarters. Panels A and B of Table XIV report the results for $N = 3, 6$, respectively. Comparing column (1) and (2), we see that adding lags substantially increases the $\bar{R}^2$. The null hypothesis of $\delta_1 = \delta_2 = 0$ is rejected at the 1% level (column 3). In both panels and across different samples, the results indicate that lags of the short rate matter for forecasting even after conditioning on today's yield curve.

The key takeaway is that short rate changes appear ex-post more predictable than the cross-sectional information in yield curve would imply. The slope decomposition suggests that the result is unlikely to be explained away with time varying risk premia, and that expectations frictions also play a role. The conclusion seems to coincide with the presence of an active central bank.

VII. Conclusions

This paper studies how agents form expectations about the short rate and thus about the future path of monetary policy actions. We show that distant lags of the yield curve forecast future short rate changes beyond information embedded in today's cross section of yields. This empirical fact leads to a broader observation: There exist systematic differences between the ex-ante real fed funds rate perceived by agents in real time and its counterpart estimated by an econometrician who works under the assumption of the FIRE.

These findings are potentially important for understanding the information content of the yield curve as a reflection of risk premia and expectations about the economy. In particular, by referring to expectations frictions, our results both support and cast light on the observation in the literature that information not contained in the current yield curve helps predict future yields and bond returns. Constructing the real rate wedge as a proxy for such expectations rigidities, we show that they induce predictable dynamics of bond returns that are distinct from the statistical and survey-based measures of bond risk premia. More generally, however, our results leave open the issue of the specific channels through which expectations frictions arise. The answer to this question matters for the interpretation of unanticipated monetary policy shocks and thus for the analysis of the real effects of monetary policy. We leave this investigation to our future research.

VIII. Figures

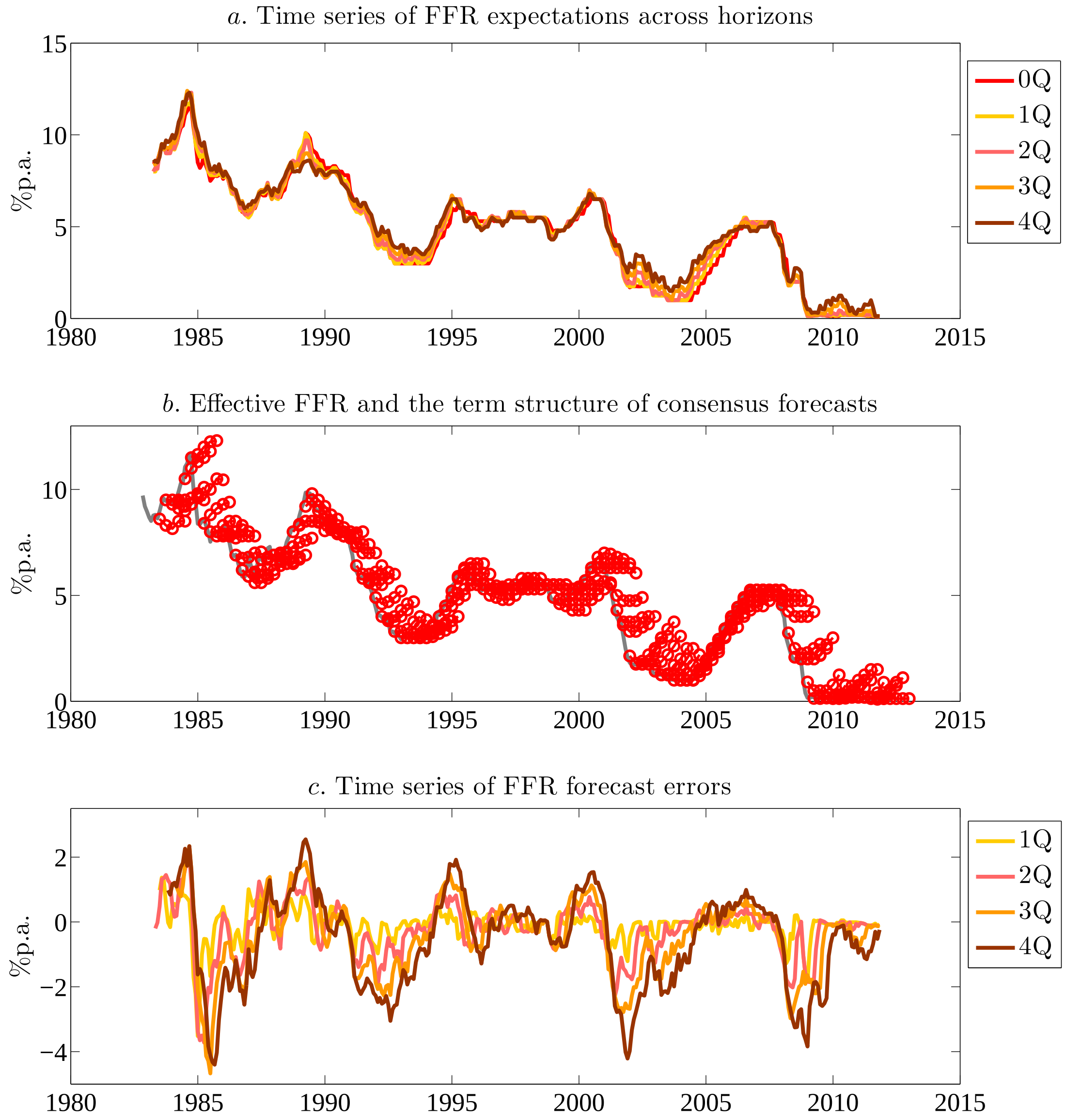


Figure 2.1: Short rate expectations

Panel *a* plots the time series of FFR forecasts from the BCFF survey at the time that the forecasts are made. The forecasts are for the current quarter up to four quarters ahead. Panel *b* plots the term structures of forecasts. For clarity, while the forecasts are given monthly, the plot shows those made in the middle of each quarter, i.e. Feb, May, Aug and Nov of each year. Panel *c* displays the time series of forecast errors for horizons from one through four quarters ahead.

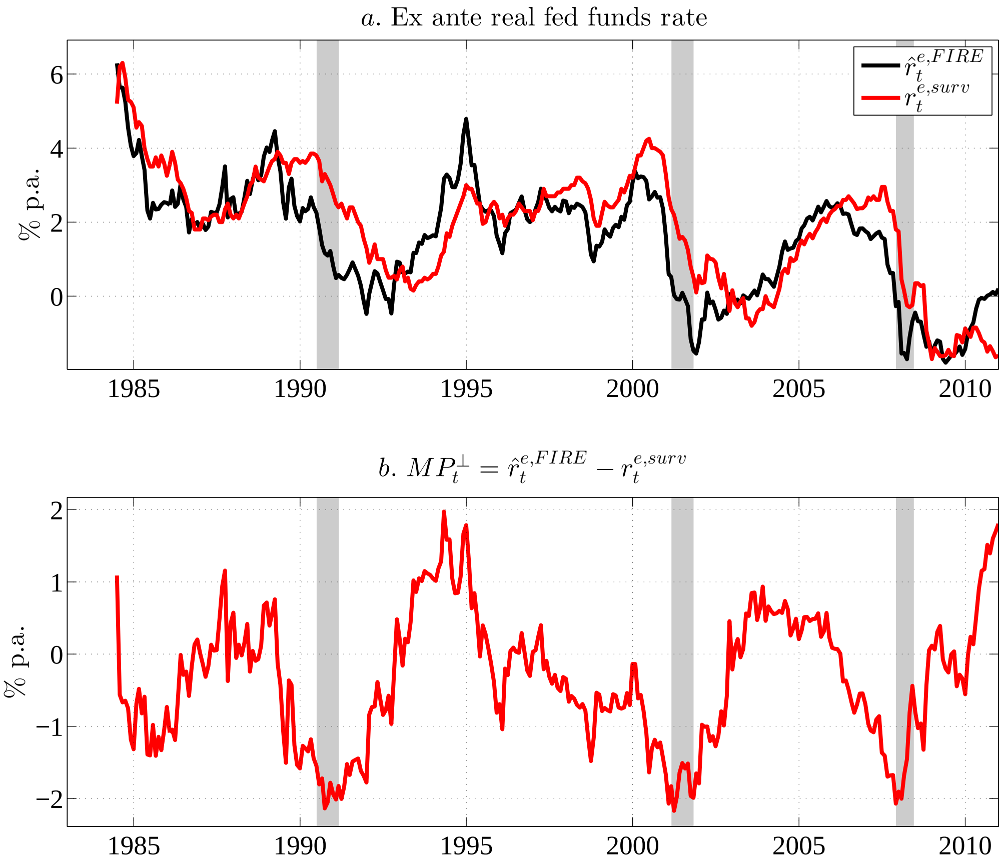


Figure 2.2: Wedge between the FIRE and survey-based real fed funds rate

Panel *a* plots the two versions of ex-ante real federal funds rate. The first one, denoted by $\hat{r}_t^{e,FIRE}$, is obtained using instrumental variables. The second, $r_t^{e,surv}$ is constructed from federal funds and inflation surveys from the BCFF survey. Panel *b* shows the dynamics of $MP_t^{\perp}$ which is obtained as a difference between $\hat{r}_t^{e,FIRE}$ and $r_t^{e,surv}$.

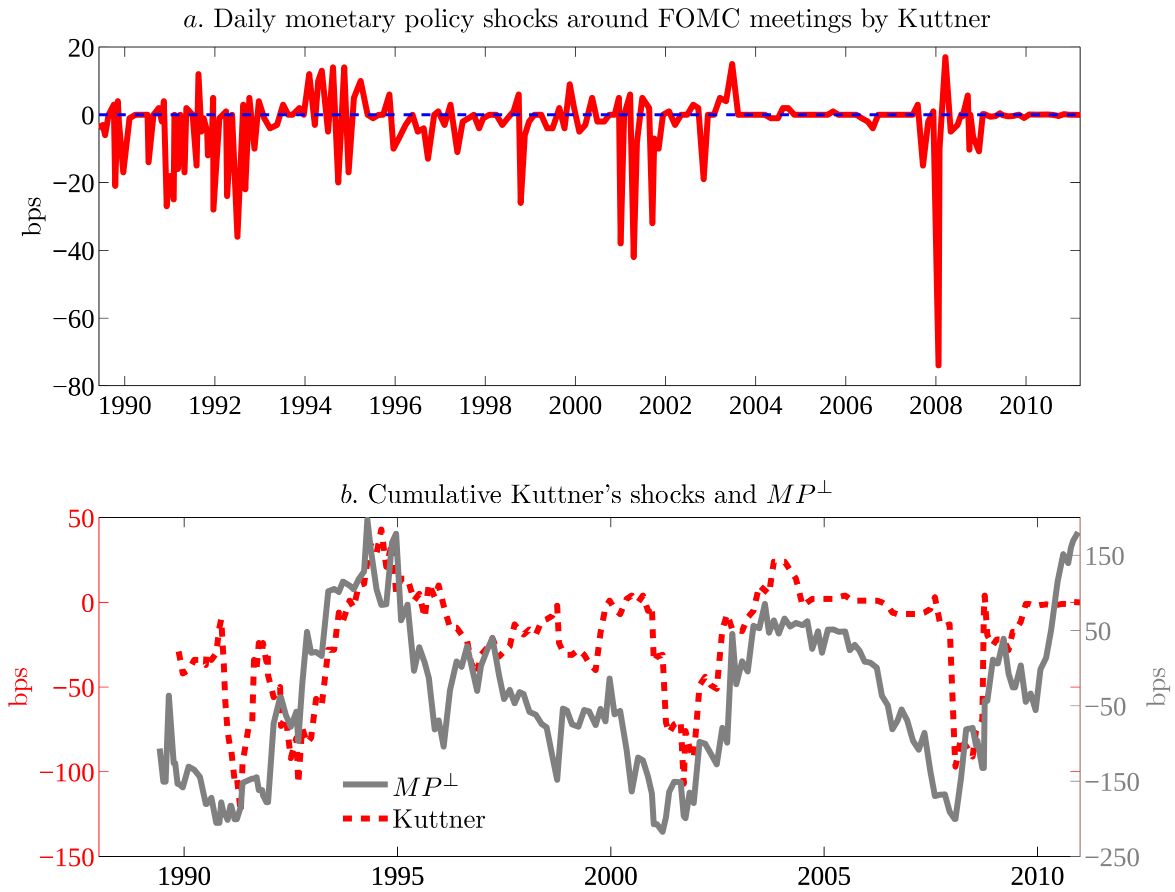


Figure 2.3: Monetary policy surprises

Panel a plots monetary policy shocks constructed by Kuttner (2001) and available on his webpage for the sample period 1989:06–2008:06, which we extend through 2010:12 using the same methodology. Kuttner obtains his measure of monetary policy surprises from daily changes in the fed fund futures around the days of target changes by the FOMC. Panel b superimposes $MP^{\perp}$ with the monetary policy surprises. Kuttner's surprises are obtained as a moving sum of the daily surprises over eight consecutive FOMC meetings. Both the left and right axis are in basis points p.a.

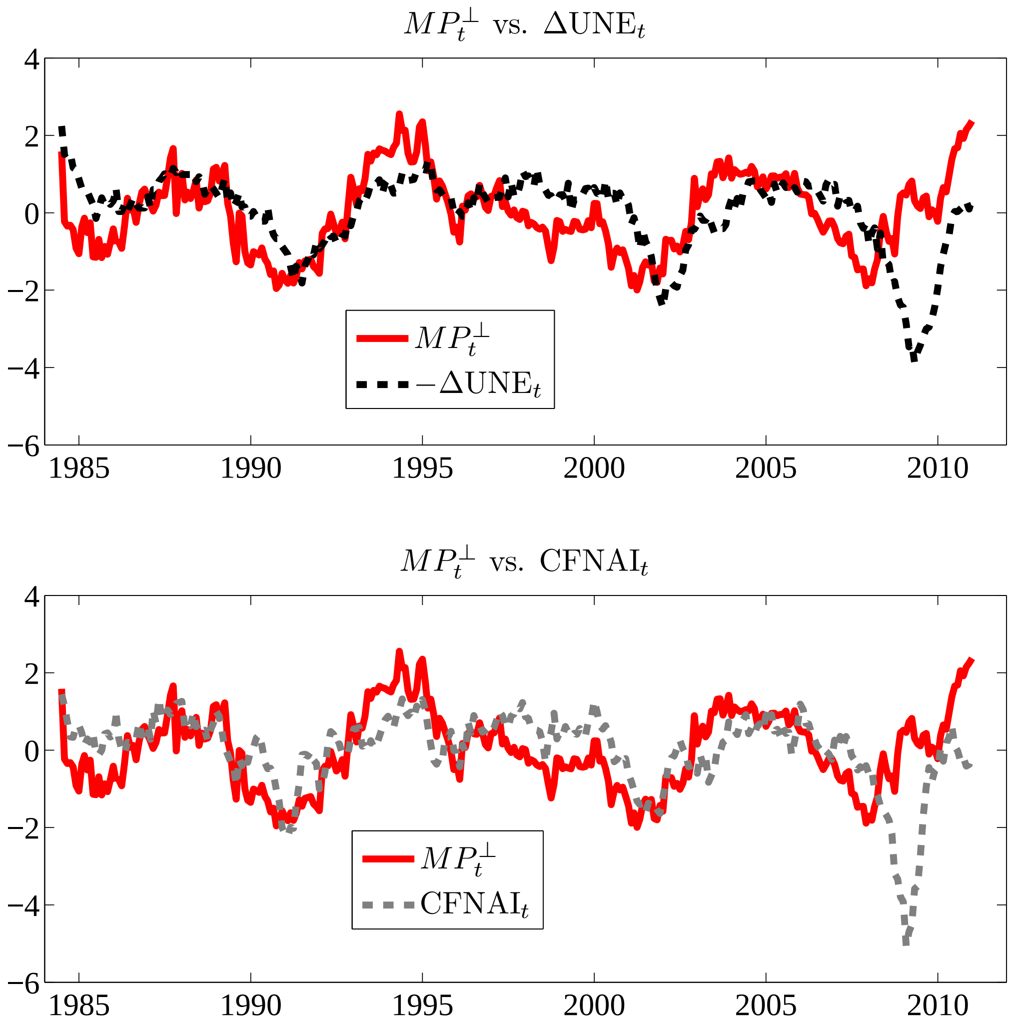


Figure 2.4: Expectation frictions and real activity

The figure plots the measure of friction in monetary policy expectations $MP_t^\perp$ and two real activity variables: the annual unemployment growth (ΔUNE$_t$) and the Chicago Fed National Activity Index (CFNAI$_t$). Unemployment growth is multiplied with -1 so that it has a positive correlation with CFNAI$_t$ and $MP_t^\perp$. For comparability, all variables are standardized.

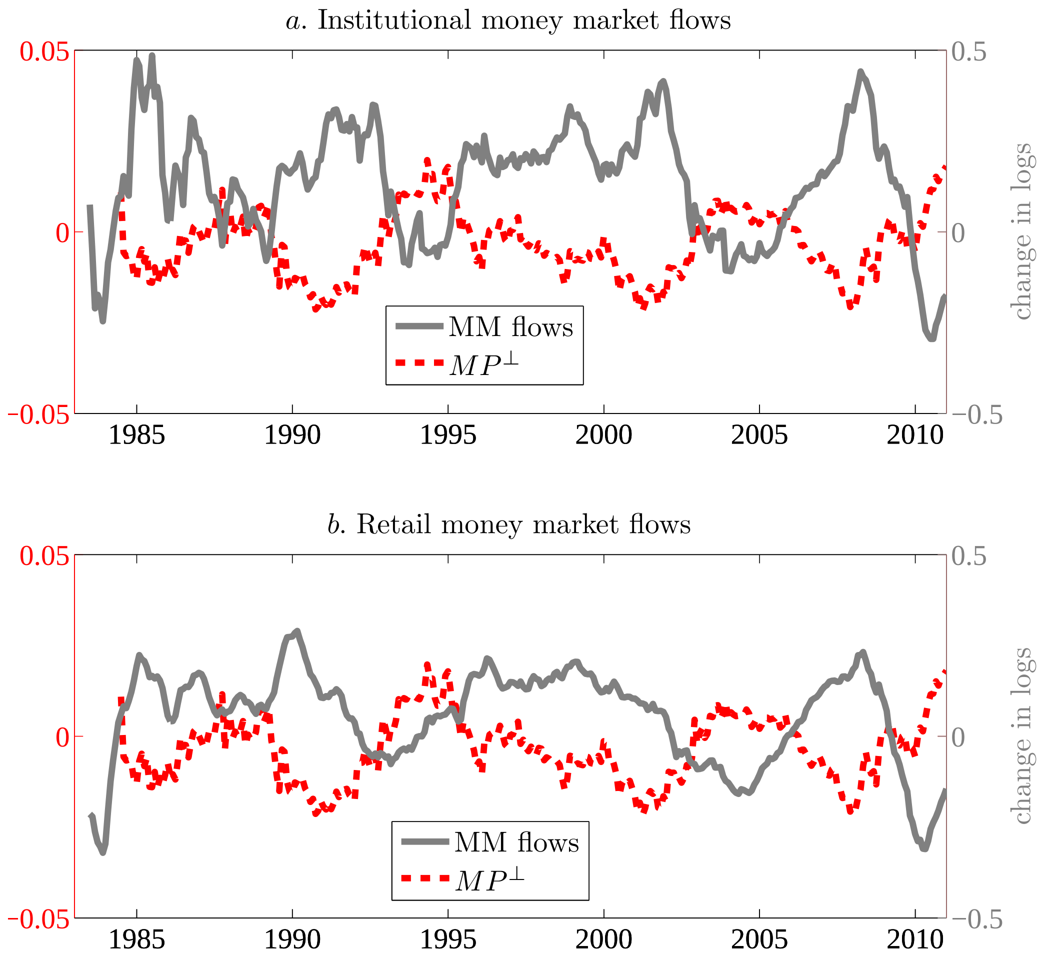


Figure 2.5: Money market flows

The figure superimposes money market flows with the contemporaneous values of $MP_t^{\perp}$. The flows are defined annual log changes in the assets of the money market funds. Panel a displays institutional flows; panel b retail flows.

IX. Tables

Table I: Private sector forecasts of short rate changes

The table reports regressions of the realized changes in the federal funds rate on the change expected by forecasters in the BCFF survey.

$$FFR_{t+h} - FFR_t = \gamma_2 + \gamma_2 \left[E_t^s(FFR_{t+h}) - FFR_t\right] + \varepsilon_{t+h}^{FE}$$

The data is monthly. T-statistics are Newey-West adjusted with 12 lags.

	$h = 1$Q	$h = 2$Q	$h = 3$Q	$h = 4$Q
γ_2	-0.09	-0.24	-0.44	-0.63
	(-1.53)	(-1.96)	(-2.27)	(-2.34)
γ_3	0.83	0.94	1.05	1.06
	(4.68)	(3.01)	(3.23)	(3.36)
t-stat ($\gamma_2 = 0$)	(-1.53)	(-1.96)	(-2.27)	(-2.34)
t-stat ($\gamma_3 = 1$)	(-0.95)	(-0.20)	(0.14)	(0.18)
$\bar{R}^2$	0.25	0.19	0.19	0.18

Table II: Private sector's expectations of monetary policy rate

The table predict forecast errors on FFR with the slope of the term structure, $S_t = y^{(20)} - y_t^{(1)}$. Panel A reports the regressions of the FFR changes on the expected change in the short rate and the lagged slope. Panel B reports analogous regressions for the forecast errors. We estimate the following equation:

$$\Delta FFR_{t,t+h} = \alpha_3 + \alpha_4 \left[E_t^s(FFR_{t+h}) - FFR_t\right] + \alpha_5 S_{t-1} + \varepsilon_{t+h} \tag{2.25}$$

and analogous for forecast errors ($FE_{t,t+h}^{FFR}$) as the LHS variable in panel B. Time subscripts and horizons are expressed as the fraction of the year, i.e. S_{t-1} is lagged by one year. The data is monthly in the period 1983–2010.

	$h = 1$Q	$h = 2$Q	$h = 3$Q	$h = 4$Q
A. Short rate changes $\Delta FFR_{t,t+h}$				
const	-0.00	-0.01	-0.01	-0.02
	(-2.27)	(-2.74)	(-3.50)	(-4.43)
$E_t^s(FFR_{t+h}) - FFR_t$	0.83	0.73	0.52	0.27
	(4.89)	(2.20)	(1.56)	(1.13)
S_{t-1}	0.06	0.20	0.41	0.66
	(1.94)	(2.33)	(3.22)	(4.95)
$\bar{R}^2$	0.32	0.29	0.33	0.37
const	-0.00	-0.01	-0.01	-0.02
	(-3.27)	(-3.78)	(-4.34)	(-4.83)
S_{t-1}	0.15	0.34	0.54	0.75
	(3.51)	(4.19)	(5.06)	(5.97)
$\bar{R}^2$	0.12	0.21	0.30	0.37
B. Forecast errors $FE_{t,t+h}^{FFR}$				
const	-0.00	-0.01	-0.01	-0.01
	(-2.23)	(-2.90)	(-3.35)	(-3.75)
S_{t-1}	0.04	0.15	0.28	0.44
	(1.52)	(2.40)	(2.95)	(3.55)
$\bar{R}^2$	0.01	0.05	0.10	0.15

Table III: Spanning of $MP_t^{\perp}$ by the cross-section of yields

This table reports the results from a regression of $MP_t^{\perp}$ on five principal components of yields. The sample period is 1984:06 through 2011:12. Standard errors are obtained with Newey-West adjustment with 15 lags.

	PC1	PC2	PC3	PC4	PC5	$\bar{R}^2$
β	-0.01	0.11	-0.12	0.17	3.54	0.19
t-stat	-1.00	1.86	-0.33	0.27	3.16	

Table IV: Forecasting annual excess bond returns

The table present the predictive regressions of realized excess bond returns across maturities. $rx^{(n)}_{t,t+1}$ is excess return with an annual holding period. The explanatory variables are two empirical measures of bond risk premium: the cycle factor of Cieslak and Povala (2011) (LHS panels) and CP factor of Cochrane and Piazzesi (2005) (RHS panels), as well as the proxy for expectations frictions $MP_t^{\perp}$. For ease of comparison, both left- and right-hand variables are standardized. The data is monthly and covers the period 1984–2010. T-statistics in parentheses in rows denoted "t(NW)" use Newey-West standard errors adjusted with 15 lags; t-statistics in brackets in rows denoted "t(H)" are based on Hodrick's reverse regressions delta method extended by Wei and Wright (2013).

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(10)}$	$rx^{(20)}$
A. $rx^{(n)}_{t,t+1} = \delta_0 + \delta_1 MP_t^{\perp} + \varepsilon_{t,t+1}$					
$MP_t^{\perp}$	-0.51	-0.45	-0.34	-0.19	-0.07
t(NW)	(-4.48)	(-3.88)	(-2.87)	(-1.57)	(-0.59)
t(H)	[-2.94]	[-2.67]	[-2.17]	[-1.55]	[-0.91]
$\bar{R}^2$	0.25	0.20	0.11	0.03	0.00

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(10)}$	$rx^{(20)}$
B1. $rx^{(n)}_{t,t+1} = \delta_0 + \delta_1 \widehat{cf}_t + \varepsilon_{t,t+1}$					
$\widehat{cf}_t$	0.49	0.54	0.63	0.72	0.70
t(NW)	(4.02)	(4.91)	(6.27)	(7.70)	(6.98)
t(H)	[2.28]	[2.70]	[3.26]	[3.88]	[4.10]
$\bar{R}^2$	0.24	0.29	0.39	0.51	0.49
B2. $rx^{(n)}_{t,t+1} = \delta_0 + \delta_1 \widehat{cf}_t + \delta_2 MP_t^{\perp} + \varepsilon_{t,t+1}$					
$\widehat{cf}_t$	0.51	0.56	0.64	0.72	0.71
t(NW)	(4.93)	(6.00)	(7.25)	(8.09)	(6.87)
t(H)	[2.68]	[3.05]	[3.51]	[4.03]	[4.11]
$MP_t^{\perp}$	-0.52	-0.47	-0.37	-0.22	-0.10
t(NW)	(-5.41)	(-4.53)	(-3.33)	(-1.89)	(-0.80)
t(H)	[-3.21]	[-2.96]	[-2.49]	[-1.87]	[-1.17]
$\bar{R}^2$	0.51	0.51	0.52	0.56	0.50

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(10)}$	$rx^{(20)}$
C1. $rx^{(n)}_{t,t+1} = \delta_0 + \delta_1 CP_t + \varepsilon_{t,t+1}$					
CP_t	0.35	0.35	0.41	0.47	0.51
t(NW)	(2.21)	(2.10)	(2.44)	(2.76)	(2.90)
t(H)	[0.68]	[1.00]	[1.47]	[1.87]	[2.01]
$\bar{R}^2$	0.12	0.12	0.17	0.22	0.26
C2. $rx^{(n)}_{t,t+1} = \delta_0 + \delta_1 CP_t + \delta_2 MP_t^{\perp} + \varepsilon_{t,t+1}$					
CP_t	0.41	0.41	0.46	0.50	0.53
t(NW)	(3.79)	(3.37)	(3.43)	(3.29)	(3.06)
t(H)	[1.32]	[1.59]	[1.97]	[2.18]	[2.10]
$MP_t^{\perp}$	-0.55	-0.49	-0.39	-0.25	-0.13
t(NW)	(-4.92)	(-4.20)	(-3.16)	(-1.90)	(-0.98)
t(H)	[-3.06]	[-2.83]	[-2.41]	[-1.86]	[-1.26]
$\bar{R}^2$	0.42	0.36	0.32	0.28	0.28

Table V: Forecasting monthly bond portfolio excess bond returns

The table present the predictive regressions of realized excess returns of bond portfolios. For instance, $rx^{(<12m)}_{t,t+1/12}$ is excess return with a monthly holding period on a portfolio of bonds whose maturities are below 12 months. Returns are in excess of the one-month Tbill rate. All returns and Tbill data is from CRSP. The explanatory variables are two empirical measures of bond risk premium: the cycle factor of Cieslak and Povala (2011) (LHS panels) and CP factor of Cochrane and Piazzesi (2005) (RHS panels), as well as the proxy for expectations frictions $MP^{\perp}_t$. For ease of comparison, both left- and right-hand variables are standardized. The data is monthly and covers the period 1984–2010. T-statistics in parentheses use Newey-West standard errors adjusted with 15 lags.

	$rx^{(<12m)}$	$rx^{(<24m)}$	$rx^{(<36m)}$	$rx^{(<60m)}$	$rx^{(<120m)}$
		A. $rx^{(n)}_{t,t+1/12} = \delta_0 + \delta_1 MP^{\perp}_t + \varepsilon_{t,t+1/12}$			
$MP^{\perp}_t$	-0.20	-0.17	-0.14	-0.08	-0.05
	(-3.51)	(-3.34)	(-2.82)	(-1.75)	(-1.07)
$\bar{R}^2$	0.04	0.03	0.02	0.00	0.00

	$rx^{(<12m)}$	$rx^{(<24m)}$	$rx^{(<36m)}$	$rx^{(<60m)}$	$rx^{(<120m)}$
B1. $rx^{(n)}_{t,t+1/12} = \delta_0 + \delta_1 \widehat{cf}_t + \varepsilon_{t,t+1/12}$					
$\widehat{cf}_t$	0.18	0.22	0.22	0.23	0.24
	(1.78)	(2.66)	(3.06)	(3.75)	(3.98)
$\bar{R}^2$	0.03	0.05	0.04	0.05	0.05
B2. $rx^{(n)}_{t,t+1/12} = \delta_0 + \delta_1 \widehat{cf}_t + \delta_2 MP^{\perp}_t + \varepsilon_{t,t+1/12}$					
$\widehat{cf}_t$	0.18	0.23	0.22	0.23	0.24
	(1.98)	(2.98)	(3.42)	(4.03)	(4.18)
$MP^{\perp}_t$	-0.21	-0.18	-0.15	-0.09	-0.06
	(-3.81)	(-3.52)	(-2.79)	(-1.71)	(-1.11)
$\bar{R}^2$	0.07	0.07	0.06	0.05	0.05

	$rx^{(<12m)}$	$rx^{(<24m)}$	$rx^{(<36m)}$	$rx^{(<60m)}$	$rx^{(<120m)}$
C1. $rx^{(n)}_{t,t+1/12} = \delta_0 + \delta_1 CP_t + \varepsilon_{t,t+1/12}$					
CP_t	0.15	0.19	0.18	0.16	0.16
	(1.54)	(2.38)	(2.40)	(2.51)	(2.50)
$\bar{R}^2$	0.02	0.03	0.03	0.02	0.02
C2. $rx^{(n)}_{t,t+1/12} = \delta_0 + \delta_1 CP_t + \delta_2 MP^{\perp}_t + \varepsilon_{t,t+1/12}$					
CP_t	0.18	0.22	0.19	0.18	0.17
	(2.08)	(3.27)	(3.29)	(3.16)	(2.91)
$MP^{\perp}_t$	-0.22	-0.19	-0.16	-0.10	-0.07
	(-3.76)	(-3.59)	(-2.95)	(-1.95)	(-1.35)
$\bar{R}^2$	0.06	0.07	0.05	0.03	0.02

Table VI: Expected returns versus expectations friction

We regress components of the realized return on a two-year bond between time t and $t+1$ on time t variables. In panel A, as dependent variables, we consider the unexpected return $rx^{(2)}_{t,t+1} - E^s_t(rx^{(2)}_{t,t+1}) = -(y^{(1)}_{t+1} - E^s_t(y^{(1)}_{t+1}))$, private sector's forecast error about the federal funds rate four quarters ahead $FE^{FFR}_{t,t+1} = FFR_{t+1} - E^s_t(FFR_{t+1})$, and the expected return component $E_t(rx^{(2)}_{t,t+1}) = f^{(2,1)}_t - E^s_t(y^{(1)}_{t+1})$. We explain the variation in the dependent variables with our measure of information rigidities $MP^{\perp}_t$, and with its components. In panel B, we regress the same dependent variables on two macro factors related to the real activity: the year-over-year growth in the rate of unemployment and the CFNAI. The data is monthly and covers the period 1987:12–2010:12; the beginning of the sample is dictated by the availability of the one-year yield forecasts in the BCFF survey.

A. Regressions of components of realized returns on $MP^{\perp}$

	Unexpected return, $rx^{(2)}_{t,t+1} - E^s_t(rx^{(2)}_{t,t+1})$						Expected return, $E^s_t(rx^{(2)}_{t,t+1})$						Forecast error, $FE^{FFR}_{t,t+1}$					
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
$MP^{\perp}_t$	-0.63						0.15						0.69					
	(-5.12)						(0.90)						(6.57)					
$r^{e,FIRE}_t$		-1.04	-0.28					0.38	0.32					1.17	0.37			
		(-4.82)	(-2.59)					(1.51)	(2.84)					(6.38)	(3.17)			
$r^{e,surv}_t$		0.95		0.11				-0.08		0.23				-1.00		-0.06		
		(4.80)		(0.77)				(-0.30)		(2.04)				(-5.74)		(-0.47)		
S_{t-1}					-0.47						-0.11						0.45	
					(-3.63)						(-0.57)						(3.57)	
$\widehat{cf}_t$						0.17						0.48						-0.10
						(1.22)						(4.33)						(-0.78)
$\bar{R}^2$	0.39	0.39	0.08	0.01	0.22	0.02	0.02	0.10	0.10	0.05	0.01	0.23	0.47	0.49	0.13	0.00	0.20	0.01

B. Regressions of components of realized returns on macro variables

	Unexpected return, $rx^{(2)}_{t,t+1} - E^s_t(rx^{(2)}_{t,t+1})$				Expected return, $E^s_t(rx^{(2)}_{t,t+1})$				Forecast error, $FE^{FFR}_{t,t+1}$			
Regressor	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
CFNAI_t	-0.39		-0.16		0.17		0.13		0.42		0.18	
	(-2.43)		(-2.20)		(1.49)		(1.55)		(2.08)		(1.92)	
ΔUNE_t		0.42		0.21		-0.05		0.01		-0.42		-0.19
		(2.63)		(2.25)		(-0.47)		(0.08)		(-2.27)		(-1.75)
$MP^{\perp}_t$			-0.56	-0.55			0.10	0.16			0.62	0.62
			(-4.29)	(-4.04)			(0.63)	(0.92)			(5.28)	(5.00)
$\bar{R}^2$	0.15	0.18	0.41	0.43	0.02	0.00	0.03	0.02	0.17	0.18	0.50	0.50

Table VII: Private sector forecasts of short rate changes

Panel A reports the root mean squared error (RMSE) in percent per annum for the out-of-sample forecasts of the FFR at horizons from one to four quarters ahead. We consider the following models: (1) the survey based forecast, (2) random walk, (3) univariate AR(2), (4) univariate AR(p) with lags selected dynamically using BIC (UDLS), (5) VAR(2) estimated recursively by OLS, (6) Bayesian VAR(2) estimated with the Minnesota prior (reset at each iteration), (7) time-varying parameters homoscedastic Bayesian VAR(2) (TVP-VAR). All models are estimated recursively with a burn-in period of 73 quarters. The data is quarterly. The out-of-sample period is 1983:Q1–2010:Q4, i.e. it coincides with the availability of survey forecasts. Panel B compares the RMSEs of forecast errors from the survey and from the fed fund futures. The sample starts in 1988:12, when the fed fund futures data become available. T-statistics test for the difference between the respective MSEs; the correlation is between the survey and futures-based forecast errors.

	$h = 1$Q	$h = 2$Q	$h = 3$Q	$h = 4$Q
A. RMSE of forecast errors (% p.a.) from different models				
(1) FFR survey	0.33	0.75	1.12	1.47
(2) RW	0.54	0.95	1.31	1.63
(3) AR(2)	0.52	0.95	1.29	1.60
(4) UDLS	0.55	0.97	1.30	1.61
(5) VAR(2) OLS	0.55	0.93	1.30	1.64
(6) VAR(2) Bayesian	1.14	1.46	1.75	2.02
(7) TVP VAR(2)	0.56	1.02	1.42	1.79
B. RMSE for surveys and fed fund futures, 1988:12-2010:12				
Fed fund futures	0.33	0.70	–	–
FFR survey	0.36	0.72	–	–
t-stat (diff = 0)	2.49	0.86	–	–
correlation	0.89	0.93	–	–

Table VIII: Correlations

The table reports unconditional correlations between forecast errors made by the private sector (professional forecasters) and the Fed staff (Greenbook). Private sector forecasts are from the BCFF survey for the FFR, and from the SPF survey for CPI inflation and unemployment. All forecasts are for four quarters ahead. The data is quarterly in the sample 1983:Q1–2006:Q4.

	$FFR^\star$	$UNE^\star$	$CPI^\star$	$FFR^\dagger$	$UNE^\dagger$	$CPI^\dagger$
$FFR^\star$	1.000					
$UNE^\star$	-0.744	1.000				
$CPI^\star$	0.317	-0.080	1.000			
$FFR^\dagger$	0.883	-0.703	0.286	1.000		
$UNE^\dagger$	-0.736	0.905	-0.121	-0.679	1.000	
$CPI^\dagger$	0.296	-0.066	0.975	0.299	-0.146	1.000

* denotes private sector (professional forecasters) forecasts; † denotes the Greenbook forecasts.

Table IX: Fed's staff monetary policy expectations from Greenbook

Panel A reports the regressions of the FFR changes on lagged expected path of the short rate and lagged monetary policy cycle. Panel B reports analogous regressions for the forecast errors. We estimate the following equation:

$$\Delta FFR_{t,t+h} = \gamma_0 + \gamma_1 \left[E_t^{GB}(FFR_{t+h}) - FFR_t\right] + \gamma_2 S_{t-1} + \varepsilon_{t+h} \tag{2.26}$$

and analogous for forecast errors ($FE^{FFR}_{t,t+h}$) as the LHS variable. Time subscripts and horizons are expressed as the fraction of the year, i.e. S_{t-1} is lagged by one year. The data is at the frequency of the FOMC meetings and spans the period 1983:3–2006:12, having 191 observations in total.

A. Short rate changes $\Delta FFR_{t,t+h}$				
	$h = 1$Q	$h = 2$Q	$h = 3$Q	$h = 4$Q
const	-0.26	-0.63	-1.02	-1.39
	(-3.04)	(-3.12)	(-3.16)	(-3.29)
$E_t^s(FFR_{t+h}) - FFR_t$	1.01	0.89	0.68	0.60
	(8.65)	(5.62)	(3.02)	(2.15)
S_{t-1}	0.08	0.22	0.40	0.56
	(2.93)	(3.39)	(3.77)	(4.05)
$\bar{R}^2$	0.37	0.31	0.28	0.30
const	-0.31	-0.63	-0.99	-1.36
	(-2.50)	(-2.68)	(-2.88)	(-3.15)
S_{t-1}	0.14	0.29	0.45	0.62
	(2.91)	(3.31)	(3.74)	(4.13)
$\bar{R}^2$	0.10	0.15	0.20	0.25
const	-0.11	-0.23	-0.30	-0.39
	(-1.87)	(-1.60)	(-1.37)	(-1.38)
$E_t^s(FFR_{t+h}) - FFR_t$	1.09	1.02	0.86	0.87
	(7.61)	(4.94)	(3.23)	(2.75)
$\bar{R}^2$	0.34	0.22	0.13	0.10
B. Forecast errors $FE^{FFR}_{t,t+h}$				
const	-0.27	-0.64	-1.04	-1.42
	(-2.67)	(-3.05)	(-3.22)	(-3.33)
S_{t-1}	0.09	0.22	0.38	0.53
	(2.64)	(3.19)	(3.48)	(3.62)
$\bar{R}^2$	0.05	0.11	0.16	0.20

Table X: IV regressions with macro expectations

The table reports the regressions of FFR forecast errors on the errors about CPI inflation and unemployment. As instruments, we use the contemporaneous oil shock and past CFNAI lagged by one quarter. Oil shock is the residual from an AR(2) estimated on the oil price change. For both instruments we report the first stage estimates. Row labeled "Weak (size, 10%)" displays the outcome of the Stock-Yogo test for the bias in standard errors. "No" indicates that we reject the null that significance level is smaller than at least 10% when the desired level is 5%, i.e. we fail to find evidence of biased standard errors due to the presence of weak instruments. T-statistics (in parentheses) use Newey-West adjustment.

	Sample: 1983:Q1–2010:Q4				Sample: 1983:Q1–2006:Q4			
Forecast errors, $FE^{FFR}_{t,t+h}$								
	$h=3Q$		$h=4Q$		$h=3Q$		$h=4Q$	
	LS	IV	LS	IV	LS	IV	LS	IV
$FE^{CPI}_{t,t+h}$	0.08	0.00	0.14	0.07	0.17	0.15	0.22	0.15
	1.36	-0.02	2.33	0.75	3.17	2.07	3.42	1.84
$FE^{UNE}_{t,t+h}$	-0.96	-0.94	-1.01	-0.96	-1.52	-1.81	-1.62	-1.99
	-3.55	-2.43	-3.76	-2.54	-8.39	-6.77	-9.73	-5.90
$\bar{R}^2$	0.39	0.37	0.47	0.45	0.54	0.52	0.61	0.58
Weak (size, 10%)	—	No	—	No	—	No	—	No
First stage								
	$FE^{CPI}_{t,t+h}$		$FE^{UNE}_{t,t+h}$		$FE^{CPI}_{t,t+h}$		$FE^{UNE}_{t,t+h}$	
	$h=3Q$	$h=4Q$	$h=3Q$	$h=4Q$	$h=3Q$	$h=4Q$	$h=3Q$	$h=4Q$
Oil shock$_{t+h}$	0.12	0.12	—	—	0.17	0.17	—	—
	5.50	5.49	—	—	6.57	6.69	—	—
CFNAI$_t$	—	—	-0.55	-0.65	—	—	-0.39	-0.46
	—	—	-4.50	-4.25	—	—	-5.10	-4.31
$\bar{R}^2$	0.29	0.27	0.37	0.34	0.22	0.21	0.22	0.21

Table XI: Tests of information frictions

Panel A, column (1) denoted "Baseline", reports estimates of equation (2.20). Columns (2)–(3) augment this regression respectively with: $MP_t^{\perp}$ in column (2), and S_{t-1} in column (3). Panels B and C perform the same test for forecast errors about unemployment and CPI inflation, respectively, i.e. forecast errors for each macro variable are regressed on the corresponding forecast update. FFR forecasts are from the BCFF survey; unemployment and CPI forecasts are from the SPF survey. The RHS variables are standardized. The data is quarterly and spans the sample period 1984:Q3–2010:Q4. T-statistics use Newey-West adjustment with 6 quarterly lags.

	A. FFR			B. Unemployment			C. CPI		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coeff.	Baseline	$MP_t^{\perp}$	S_{t-1}	Baseline	$MP_t^{\perp}$	S_{t-1}	Baseline	$MP_{t-1}^{\perp}$	S_{t-1}
Horizon, $h = 1Q$									
β_0	-0.04	-0.09	-0.09	-0.04	-0.04	-0.04	-0.05	-0.05	-0.05
	(-1.03)	(-2.85)	(-2.57)	(-1.03)	(-1.12)	(-1.14)	(-0.69)	(-0.69)	(-0.70)
β_1	0.28	0.09	0.14	0.15	0.13	0.14	0.22	0.21	0.22
	(6.54)	(2.57)	(4.77)	(2.55)	(1.96)	(2.53)	(3.01)	(2.88)	(2.99)
β_X		0.16	0.07		-0.06	-0.06		0.05	0.03
		(3.99)	(1.88)		(-1.89)	(-2.03)		(0.63)	(0.26)
$\bar{R}^2$	0.14	0.26	0.16	0.20	0.23	0.23	0.03	0.03	0.02
Horizon, $h = 2Q$									
β_0	-0.14	-0.24	-0.24	-0.02	-0.02	-0.02	-0.13	-0.13	-0.13
	(-1.39)	(-3.45)	(-2.86)	(-0.25)	(-0.28)	(-0.27)	(-1.05)	(-1.06)	(-1.06)
β_1	0.46	0.13	0.25	0.24	0.21	0.22	-0.01	-0.02	-0.01
	(4.92)	(1.75)	(3.67)	(3.22)	(2.24)	(3.09)	(-0.09)	(-0.17)	(-0.06)
β_X		0.38	0.18		-0.13	-0.13		0.08	0.06
		(5.24)	(2.56)		(-4.29)	(-2.26)		(0.76)	(0.43)
$\bar{R}^2$	0.13	0.34	0.18	0.23	0.29	0.29	0.00	0.00	0.00
Horizon, $h = 3Q$									
β_0	-0.26	-0.41	-0.41	0.02	0.02	0.02	-0.22	-0.22	-0.22
	(-1.71)	(-4.06)	(-3.17)	(0.22)	(0.25)	(0.24)	(-1.66)	(-1.67)	(-1.68)
β_1	0.75	0.21	0.39	0.30	0.24	0.25	0.01	-0.01	0.02
	(5.17)	(2.74)	(4.10)	(3.21)	(1.96)	(2.75)	(0.15)	(-0.11)	(0.19)
β_X		0.62	0.32		-0.25	-0.24		0.13	0.13
		(6.35)	(3.19)		(-5.79)	(-2.52)		(1.04)	(1.16)
$\bar{R}^2$	0.17	0.46	0.25	0.17	0.29	0.28	0.00	0.00	0.00

Table XII: Evidence from money market flows

Panel A runs predictive regressions of changes in the FFR and forecast errors on the annual money market flows, for retail and institutional money market funds, respectively. Panel B explains the annual change in the flows with two variables: contemporaneous measure of expectations friction $MP^{\perp}_{t+1}$, and slope S_t. In column (2) we report the results using $MP^{\perp}$ and the control variables for the flight to quality and liquidity: Pastor-Stambaugh market-wide liquidity, Hu-Pan-Wang noise illiquidity, Fontaine-Garcia value of funding liquidity, and stock market volatility VXO. Annual flows are log year-over-year changes in the money market funds. All variables are standardized. T-statistics are Newey-West adjusted with 15 lags. The sample is monthly in the period 1984:6-2010:12, except for the regression in panel B marked with * where we use the period 1987:01–2010:12, due to availability of some of the controls.

A. Predicting short rate with flows: $Y_{t,t+1} = \alpha + \beta \text{Flow}_{t-1,t} + \varepsilon_{t,t+1}$

	Retail flows		Institutional flows	
Dependent $Y_{t,t+1}$:	$\Delta FFR_{t,t+1}$	$FE^{FFR}_{t,t+1}$	$\Delta FFR_{t,t+1}$	$FE^{FFR}_{t,t+1}$
	(1)	(2)	(1)	(2)
β	-0.38	-0.17	-0.52	-0.44
	(-3.17)	(-1.48)	(-4.35)	(-4.25)
$\bar{R}^2$	0.14	0.03	0.27	0.19

B. Predicting flows: $\text{Flow}_{t,t+1} = \alpha + \beta X_t + \varepsilon_{t,t+1}$

	Retail flows			Institutional flows		
Regressor X_t:	$MP^{\perp}_{t+1}$	$MP^{\perp}_{t+1}$ w/controls*	S_t	$MP^{\perp}_{t+1}$	$MP^{\perp}_{t+1}$ w/controls*	S_t
	(1)	(2)	(3)	(1)	(2)	(3)
β	-0.51	-0.77	-0.77	-0.75	-0.80	-0.72
	(-3.79)	(-5.97)	(-6.62)	(-8.11)	(-6.72)	(-6.58)
$\bar{R}^2$	0.25	0.52	0.59	0.56	0.62	0.52

Table XIII: Statistical model of the short rate with lags, 1875-2011

Panel A reports the results from the predictive regression of one-year change in the short rate. Four lags of the short rate are used as predictors. Maximum lag is 16 quarters. Panel B of the table reports the decomposition of the slope given by (2.23). Panel C reports the estimation results for a predictive regression using the fitted value (short rate expectations) and the residual (risk premia) from (2.23). Excess return on a ten-year bond is constructed from total return indices on ten-year bond and three-month Tbill, both obtained from the GFD database. Numbers in brackets [·] in the $\bar{R}^2$ column report the $\bar{R}^2$'s obtained with the Cochrane-Piazzesi factor. Panel D reports the results for predicting the short rate changes using $\hat{S}_t$. Data are quarterly and are from the Global Financial Database. The slope is constructed as a difference between ten-year par coupon yield on Treasuries and the three-month yield. The three-month yield is spliced from the three-month commercial paper (1875-1933) and three-month Treasury bill (1934-2011). Newey-West adjusted (6 lags) t-statistics are reported in parentheses.

A. $\Delta i_{t,t+1} = \alpha_c + \sum_{j\in 0,1,2,4} \alpha_j i_{t-j} + \varepsilon^i_{t+1}$

Sample	α_0	α_1	α_2	α_4	$\bar{R}^2$	$\bar{R}^2$ no lags
1875:3-1913:4	-0.81	0.02	0.10	-0.05	0.40	0.40
	(-5.21)	(0.14)	(1.10)	(-1.54)		
1914:1-1951:2	-0.22	0.00	0.15	-0.03	0.07	0.05
	(-1.81)	(0.02)	(1.27)	(-0.20)		
1951:3-1979:2	-0.52	-0.37	0.34	0.51	0.39	0.04
	(-3.47)	(-2.78)	(3.26)	(3.71)		
1984:1-2011:4	0.02	-0.47	-0.01	0.23	0.41	0.11
	(0.27)	(-3.63)	(-0.10)	(3.16)		
1951:3-2011:4	-0.01	-0.27	0.04	0.14	0.13	0.06
	(-0.10)	(-1.88)	(0.38)	(1.59)		

B. $S_t = \delta_c + \sum_{j\in 0,1,2,4} \delta_j i_{t-j} + RP^S_t$

Sample	δ_0	δ_1	δ_2	δ_4	$\bar{R}^2$	$\bar{R}^2$ no lags
1875:3-1913:4	-1.02	-0.01	0.04	0.08	0.96	0.94
	(-63.31)	(-0.56)	(2.64)	(4.23)		
1914:1-1951:2	-0.78	0.02	0.10	0.08	0.94	0.91
	(-12.20)	(0.29)	(2.91)	(1.65)		
1951:3-1979:2	-0.56	0.14	0.31	0.19	0.79	0.10
	(-9.96)	(3.16)	(5.71)	(3.05)		
1984:1-2011:4	-0.39	-0.10	0.22	0.23	0.68	0.15
	(-5.71)	(-1.42)	(3.60)	(7.45)		
1951:3-2011:4	-0.44	0.08	0.21	0.20	0.70	0.12
	(-9.79)	(1.46)	(4.33)	(5.15)		

	C. $rx^{(10)}_{t+1} = \beta_0 + \beta^e \hat{S}_t + \beta^{rx} \widehat{RP}^S_t + \varepsilon^{rx}_{t+1}$			D. $\Delta i_{t,t+1} = \beta_0 + \beta^e \hat{S}_t + \varepsilon^e_{t+1}$		
Sample	β^e	β^{rx}	$\bar{R}^2$	β^e	$\bar{R}^2$	$\bar{R}^2$ slope
1875:3-1913:4	0.09	0.22	0.05	1.13	0.39	0.35
	(0.75)	(0.92)		(5.70)		
1914:1-1951:2	0.15	0.40	0.18	0.28	0.07	0.06
	(0.98)	(2.37)		(1.60)		
1951:3-1979:2	-0.03	0.31	0.08 [0.04]	0.56	0.17	0.14
	(-0.25)	(2.59)		(3.61)		
1984:1-2011:4	-0.03	0.56	0.30 [0.21]	0.50	0.12	0.09
	(-0.22)	(5.68)		(2.65)		
1951:3-2011:4	0.16	0.38	0.16 [0.20]	0.26	0.02	0.03
	(1.21)	(3.90)		(1.25)		

Table XIV: Predicting short rate changes with lags, 1952-2011

The table reports the results from a predictive regression for one-year change in the short term interest rate. We use principal components and lags of the short term interest rate. Panel A reports the results for three principal components and Panel B for six principal components. Each panel displays the results for three different sub-samples; (*i*) pre-Volcker period and (*ii*) post-Volcker period and (*iii*) full sample. Lags of the short rate are optimally selected using BIC selection criteria. The maximum lag is 16 quarters. Selected lags are reported in the last two columns of each panel. First column reports the adj. R^2 obtained with PCs, the second column reports the adj. R^2 using PCs and two optimally selected lags. The data are quarterly and PCs are obtained from the Fama-Bliss yield data augmented by the three-month Tbill to capture the information at the short end.

	$\bar{R}^2$ PCs only	$\bar{R}^2$ PCs & lags	Wald p-val	Opt. lag1 (Qtrs.)	Opt. lag2 (Qtrs.)
	(1)	(2)	(3)	(4)	(5)
	A. $i_{t+1} - i_t = \beta_0 + \sum_{j=1}^{3} \beta_j PC_t^j + \delta_1 i_{t-lag1/4} + \delta_2 i_{t-lag2/4} + \varepsilon_{t+1}$				
1952:3–1979:2	0.09	0.39	0.00	9	14
1984:1–2011:4	0.38	0.53	0.00	4	16
1952:3–2011:4	0.14	0.21	0.01	3	16
	B. $i_{t+1} - i_t = \beta_0 + \sum_{j=1}^{6} \beta_j PC_t^j + \delta_1 i_{t-lag1/4} + \delta_2 i_{t-lag2/4} + \varepsilon_{t+1}$				
1952:3–1979:2	0.08	0.36	0.00	13	–
1984:1–2011:4	0.37	0.55	0.00	5	16
1952:3–2011:4	0.18	0.25	0.01	3	16

X. Appendix

X.A. Lag selection for the short rate

Table XV: The predictive power of lags for short rate changes

The table reports the predictive power of the short rate i_t for the future one-year change in the short rate $i_{t+1} - i_t$, considering the short rate at different lags k. We use the BIC to select the lag length for the short rate considering all possible combinations of lags from zero to 16. For the best specification, the table reports the BIC, the $\bar{R}^2$, the specific lags chosen, and the relative probability (Rel.prob. no lags). Rel.prob. no lags measures the probability of a model without lags, i.e. based on i_{t-k} $(k = 0)$, relative to the first best specification with lags. Relative probability is obtained as: $\exp\{(BIC_{\text{best}} - BIC_{k=0})T/2\}$, where $BIC = \ln(\widehat{\sigma^2}) + \ln(T)n/T$, n is the number of regressors, $\widehat{\sigma^2} = SSE/T$ of the regression, and T is the sample size. Similarly, Rel. prob. fix lags measures the probability of the model with fixed lags relative to the best lag specification. The data is quarterly.

$$i_{t+1} - i_t = \gamma_c + \textstyle\sum_k \gamma_k i_{t-k} + \varepsilon_{t+1}$$

Sample	BIC	$\bar{R}^2$	$\bar{R}^2$ no lags	Rel. prob. no lags	$\bar{R}^2$ fixed lags	Rel. prob. fix lags	Lag 1	Lag 2	Lag 3
1875:3–1913:4	0.53	0.42	0.37	0.05	0.38	0.00	0	8	14
1914:1–1951:2	0.00	0.15	0.06	0.02	0.07	0.00	1	7	11
1951:3–1979:2	0.33	0.31	0.00	0.00	0.30	0.01	2	14	–
1984:1–2011:4	0.27	0.42	0.11	0.00	0.42	0.02	4	16	–
1951:3–2011:4	0.86	0.15	0.06	0.00	0.13	0.00	3	16	–

X.B. Survey data

To test for potential biases in the FFR forecasts, we regress future $t+h$ realizations of the FFR on the time-t forecasts, for h ranging from one to four quarters ahead, $FFR_{t+h} = \alpha + \beta E^s_t(FFR_{t+h}) + \varepsilon_{t,t+h}$. An unbiased forecast implies that $\alpha = 0$ and $\beta = 1$ (Mincer and Zarnowitz, 1969). Table XVI summarizes the results. We fail to reject the null at all horizons suggesting the private sector reports unbiased forecasts of the future FFR.

Do survey forecasts match the yield curve dynamics?

We test whether FFR forecasts are a good approximation to the market-wide consensus about the path of the short rate that is reflected in the yield curve. A yield on a zero coupon bond is a sum of the average short rate that is expected to prevail until the maturity of the bond and a risk premium. Therefore, we can decompose one-year nominal yield $y^{(1)}_t$ into short rate expectations and risk premia by averaging the available FFR forecasts over the current quarter through four quarters ahead:

Table XVI: Testing for survey bias

Table reports the Mincer-Zarnowitz test for survey bias for four forecasting horizons: one ($h = 1Q$) through four ($h = 4Q$) quarters. The joint null hypothesis is $\alpha = 0, \beta = 1$. The standard errors are obtained by Newey-West adjustment with 12 lags.

$$FFR_{t+h} = \alpha + \beta E_t^s(FFR_{t+h}) + \varepsilon_{t,t+h}$$

	h=1Q	h=2Q	h=3Q	h=4Q
α	-0.14	-0.28	-0.47	-0.52
	(-1.90)	(-1.54)	(-1.47)	(-1.07)
β	1.01	1.01	1.01	0.99
	(62.84)	(27.39)	(16.09)	(10.47)
pval ($\beta = 1$)	0.25	0.40	0.42	0.54
$\bar{R}^2$	0.98	0.92	0.82	0.68

Table XVII: Forecast errors across monetary policy regimes

The table reports the means and standard deviations of the forecast errors across forecast horizons from one to four quarters ahead. We condition on the monetary policy regime: easing, tightening and neutral. The regimes are identified on a daily frequency using changes in the FFR target: easing (tightening) episode is defined as the time from the day on which the target FFR has increased (decreased) to the next monetary policy move. Neutral regime is when there has been no monetary policy action for longer than the span between two FOMC meetings. We identify 75 months as tightening, 94 months as easing and 140 months as neutral. From the daily data we construct the end of month series.

	h=Q1	h=Q2	h=Q3	h=Q4
Tightening, $N = 75$ months				
mean (μ_T)	0.18	0.40	0.49	0.60
std	0.32	0.56	0.78	0.88
Easing, $N = 94$ months				
mean (μ_E)	-0.32	-0.77	-1.14	-1.43
std	0.51	0.73	1.04	1.37
Neutral, $N = 140$ months				
mean (μ_N)	-0.09	-0.23	-0.41	-0.62
std	0.19	0.47	0.76	1.13
Z-test ($\mu_E = \mu_T$)	3.94	9.15	12.78	15.89
pval	0.00	0.00	0.00	0.00

$$y_t^{(1)} = \underbrace{\gamma_0}_{-6e^{-4}\ [-0.74]} + \underbrace{\gamma_1}_{0.99\ [62.62]} \frac{1}{5}\sum_{k=0}^{4} E_t^s(FFR_{t+\frac{k}{4}}) + \nu_t, \quad \bar{R}^2 = 0.99, \tag{2.27}$$

where $E_t^s(FFR_{t+h})$ denotes the time-t survey-based forecast of the FFR at horizon h (expressed in years). T-statistics (in brackets) are Newey-West adjusted with 12 monthly lags. Note that the regression jointly tests the accuracy of survey data and decomposes $y_t^{(1)}$ into short rate expectations and risk premia comprised in ν_t. Hence, $\nu_t = RP_t + \gamma_1 \epsilon_t$ where ϵ_t represents the survey inaccuracies, and RP_t measures the variation in the risk premium. The estimates suggest that the median survey responses at different horizons quite accurately represent market expectations about the future path of the monetary policy, as we cannot reject the hypothesis that $\gamma_0 = 0$ and $\gamma_1 = 1$ at the standard significance levels. Moreover, since expectations explain nearly all variation in the one-year yield, the risk compensation and/or survey inaccuracies can be assumed to be small.

Table XVIII: Forecast errors for unemployment and CPI inflation

Table reports RMSE for unemployment and CPI inflation. The level of variables and the RMSE is in percentages. We report the RMSE for the Greenbook and SPF forecasts. The numbers in parentheses are RMSEs divided by the standard deviation of the realized inflation and unemployment, respectively. The data is quarterly. In Panel I, the sample period is 1983:Q2–2006:Q4, i.e. when the Greenbook sample ends. In Panel II, the sample period is 1974:Q4–1991:Q4 for unemployment and 1981:Q3–1991:Q4 for CPI inflation. The end of the sample used in Panel II is consistent with the end of sample used by Romer and Romer (2000).

	h=Q1	h=Q2	h=Q3	h=Q4
Panel I. 1983:Q1–2006:Q4 sample				
A. Greenbook				
UNEMPL	0.30	0.42	0.56	0.67
	(0.23)	(0.33)	(0.44)	(0.52)
CPI	1.60	1.72	1.76	1.81
	(0.92)	(0.99)	(1.02)	(1.04)
B. SPF				
UNEMPL	0.28	0.42	0.56	0.66
	(0.22)	(0.32)	(0.43)	(0.51)
CPI	1.56	1.64	1.70	1.77
	(0.90)	(0.95)	(0.98)	(1.02)
Panel II. Pre-1992 sample				
A. Greenbook				
UNEMPL	0.43	0.63	0.74	0.87
(1974:Q4–1991:Q4)	(0.43)	(0.64)	(0.75)	(0.88)
CPI	2.10	2.19	2.03	2.15
(1981:Q3–1991:Q4)	(1.21)	(1.26)	(1.17)	(1.24)
B. SPF				
UNEMPL	0.44	0.64	0.79	0.95
(1974:Q4–1991:Q4)	(0.45)	(0.65)	(0.80)	(0.97)
CPI	1.94	2.13	2.18	2.34
(1981:Q3–1991:Q4)	(1.12)	(1.23)	(1.25)	(1.35)

Chapter 3

Information in the term structure of yield curve volatility

Unless held till maturity,[1] Treasury bonds are risky investments, and their riskiness changes over time and with the maturity of the bond. As one stark example, the current interest rate environment in the US suggests that we can learn a great deal about macroeconomic and financial conditions by looking into bond volatility. With short-term interest rates bound close to zero for almost two years and expected to remain low in the near future, long-term yields have become increasingly autonomous. Indeed, a clear near-term outlook for the monetary policy has tamed the short-end of the yield curve, but at the same time it has raised new questions about the long-run policy impact on the economy, thus contributing to increased volatility at longer maturities.

The US Treasury market is the largest and the most liquid debt market in the world. It provides investors, central banks and governments worldwide not only with the store of value and liquidity but Treasuries also serve as the main source of collateral in repurchase

[1]This chapter is based on the paper with the identical title written in collaboration with Anna Cieslak from the Kellogg School of Management, Northwestern University. We thank Torben Andersen, Luca Benzoni, Snehal Banerjee, John Cochrane, Jerome Detemple, Paul Söderlind, Viktor Todorov, Fabio Trojani, Pietro Veronesi, Liuren Wu, Haoxiang Zhu and seminar and conference participants at the New York University (Stern), Federal Reserve Bank of New York, Federal Reserve Bank of Chicago, Bank of Canada, Baruch College, University of St. Gallen, 4th Financial Long-Run Risk Forum (Institute Luis Bachelier), European Finance Association Meeting (2010), 3rd Annual SoFiE Meeting, EC^2, European Winter Finance Summit (Skinance), University of Chicago PhD brownbag, TADC London Business School, SFI NCCR PhD Workshop in Gerzensee for their comments. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK project under the direction of Fabio Trojani).

and OTC derivatives transactions. The fluctuations in Treasury yields and in their volatility have substantial impact on a wide range of asset markets. A rise in interest rate volatility may signal macroeconomic uncertainty and trigger a monetary policy action, but might also lead to an increase in haircuts for Treasury bonds used as collateral in about half of the repo transactions, and thus diminish the lending capacity in the financial system.[2] As such, the today's environment makes the understanding of interest rate volatility especially compelling.

Using comprehensive data on transactions in the US Treasury market, we explore the key drivers of volatility in government bonds. To this end, we decompose the stochastic covariance matrix of on-the-run yields into two volatility components corresponding to yield factors with a short and long duration, respectively, and an additional variable that captures the degree of comovement between the short and long maturity segment of the yield curve. This decomposition serves to uncover the following results.

First, we ask what interest rate volatility can tell us about the funding conditions in the market. We find an intriguing lead-lag relationship between yield volatility and market-wide liquidity. The long-duration volatility component captures almost half of the variation in a broad measure of the value of funding liquidity. That element of liquidity is visible in the on-/off-the-run premium across bond maturities and reflects the tightness of financial conditions in asset markets (Fontaine and Garcia, 2011). Typically, it also preempts an action of the monetary authority. An increase in the long-duration volatility component predicts a decline in the value of funding liquidity up to six months ahead with an R^2 of nearly 50%. The short-duration volatility, instead, dominates in explaining the transitory episodes of liquidity dry-ups that do not necessarily trigger a monetary policy reaction (e.g. GM/Ford downgrade), but reflect the amount of arbitrage capital available in the market (Hu, Pan, and Wang, 2013). Importantly, an increase in the short-end volatility component predicts that the arbitrage capital will be scarce in the future.

Second, we explore the macroeconomic drivers of volatility in yields. A combination of expectations and uncertainty about monetary policy, inflation and real activity accounts for

[2]According to SIFMA quarterly report, in the third quarter of 2010 the outstanding volume of repos and reverse repos among primary dealers in US government securities exceeded $4.62 trillion on average per day (SIFMA, 2010). The argument that increased volatility of Treasuries can lead to diminished lending capacity in the system has gained traction recently among practitioners as the outlook for the US economy has become more uncertain and the government debt remains at historically high levels (see e.g. JP Morgan, 2011).

up to 30% of the variation in yield volatility. Short- and long-run volatility are related to different macro variables. The long-end volatility shows a pronounced response to the real activity measures such as the uncertainty about unemployment, and to expectations about the path of monetary policy. Quite differently, the short-term volatility is most strongly linked to the monetary policy uncertainty. Finally, the comovement term is associated with the uncertainty about inflation, monetary policy and the real economy, showing that those variables influence the extent of correlation between yields across maturities.

Third, we study how interest rate volatility is related to the risk compensation required by investors for holding Treasury bonds. When short-duration volatility is high, as induced by an increased monetary policy uncertainty, bond risk premia tend to rise. This effect is statistically and economically strong for bonds with a two-year maturity, but it dissipates quickly as the maturity increases. Given that, as we show, the amount of interest rate volatility has an imperceptible effect on the cross section of yields, our evidence contributes to the current literature that studies "hidden" factors in bond risk premia (Duffee, 2011; Joslin, Priebsch, and Singleton, 2010).

Two aspects of our approach are new to the literature and instrumental in obtaining the above results. On the methodological front, we are the first to embed information from the realized covariance matrix of yields identified with high-frequency data within an estimation of a new term structure model. High-frequency sampling of the Treasury zero curve allows us to extract factors in on-the-run yields that are invisible at lower sampling frequencies but carry information about liquidity, uncertainty and risk premia. The term structure model with a stochastic covariance matrix of yields, in turn, imposes an economically intuitive structure on the data, and allows us to perform the decomposition of volatilities in a way that ensures their consistency with yield curve factors themselves.

Our focus on the Treasury market provides a different perspective of interest rate volatility than the Libor implied volatility in swaptions, caps and floors.[3] Most importantly, with help of high frequency data on cash bonds, we obtain closest possible measurement of second moments of yields that we model and aim to understand. Other important differences can also be traced back to the defining features of the spot and derivatives markets. Due to their liquidity and safety features, Treasury bonds are held by a wide range of investors.

[3]Interest rate derivatives have been the focus of many authors (e.g. Joslin, 2010; Trolle and Schwartz, 2009, 2011).

They are also directly linked to the macroeconomic outlook through the monetary policy channel. In contrast, Libor derivatives are traded by a narrower investor base, and reflect specific effects such as the hedging demand induced by the corporate bond issuance or mortgage convexity hedging (Azarias, 2010; Duarte, 2008). Hence, their link to liquidity and macroeconomy can be less direct. For robustness, we show that the implied volatility from short-maturity Treasury options, the Merrill Option Volatility Estimate (MOVE), is well explained by the trailing realized yield volatility. However, many of our findings are hard to establish using option-based volatility estimates because such estimates confound into a single variable information from securities with very different durations.

Related literature

Our paper is related to the literature on realized volatility, term structure modeling with stochastic volatility, and the research into the role of US Treasury bonds in liquidity and collateral provision.

Recent advances in high-frequency econometrics have encouraged a model-free look into the statistical properties of bond volatility. Andersen and Benzoni (2010) test empirically the volatility implications of affine term structure models (ATSMs) using realized volatility over the 1991–2000 period obtained from the GovPx database. In line with the early evidence in Collin-Dufresne and Goldstein (2002), they confirm that systematic volatility factors are largely independent from the cross section of yields, and call for extensions of the popular models in the volatility dimension.[4]

In the latent factor domain, several recent papers document that benchmark models with stochastic volatility face difficulties in explaining variation in yield volatility, and those difficulties become increasingly severe at the long end of the yield curve. Collin-Dufresne, Goldstein, and Jones (2009, CDGJ) report that over the 1988–2005 sample variance series generated by a standard three-factor model, $A_1(3)$, are essentially unrelated to the model-free conditional volatility measures. Using the same model, Jacobs and Karoui (2009) find a correlation between model-implied and EGARCH volatility reaching up to 75% over the 1970–2003 period. However, in the more recent 1991–2003 sample this correlation breaks

[4]Several other 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).

down and becomes negative at the long end of the curve. Recently, in a study of Japanese yields in the zero bound environment, Kim and Singleton (2011) arrive at a similar conclusion based on several different models. These authors suggest that additional volatility factors may be needed to explain the volatility at the long maturity range. We contribute to this discussion by designing a new term structure model that accommodates the multivariate nature of yield volatilities, and support its estimation with the information from the realized covariance matrix of yields. Explaining 98% of volatility both at the short and long end of the yield curve in the 1992–2010 period, we find that these two elements of our approach alleviate the problems documented in the literature.

Andersen, Bollerslev, Diebold, and Vega (2007) and Jones, Lamont, and Lumsdaine (1998) show that relative to other liquid asset markets, bond prices tend to provide a clear and pronounced reaction to economic news. These studies suggest that a rich economic content is present in bond volatilities. The model model-based decomposition of the volatility curve allows us to study how its components react to measures of economic and financial conditions.

Several recent studies introduce stochastic volatility into macro-finance term structure models.[5] Adrian and Wu (2009) and Campbell, Sunderam, and Viceira (2012) highlight the importance of a stochastic covariance between the real pricing kernel and (expected) inflation in determining excess bond returns. These models attach economic labels to different yield volatility components. To the extent that the volatility itself remains unobservable or is extracted from an auxiliary model, the identification of its components relies on specific assumptions. Explaining the volatility curve per se is not in direct focus of those models. We, instead, start from a latent factor model, and having decomposed yields and volatilities, we aim to understand the impact of economic and financial quantities on both curves.

Given the role that US Treasuries play in the current financial system, it is important to understand how volatility of bonds relates to the fluctuations in demand for liquidity and collateral. A growing literature focuses on extracting information about liquidity premia from the cross section of bonds with different age (e.g. Vayanos and Weil, 2008; Fontaine and Garcia, 2011). Recently, Hu, Pan, and Wang (2013) compose a noise illiquidity measure as an average yield pricing error of Treasury bonds. We show that yield volatility of short and long duration exhibits distinct links to those concepts of liquidity. We conclude that

[5] Examples of such models are Campbell, Sunderam, and Viceira (2012), Adrian and Wu (2009), Hautsch and Ou (2008), Bekker and Bouwman (2009) or Haubrich, Pennacchi, and Ritchken (2011).

observing volatility in on-the-run Treasuries can inform policy makers about the funding conditions in the market.

I. Data

We use high-frequency nominal Treasury bond data spanning two long expansions, two recessions and three monetary cycles in the US economy. This section describes our data set. Technical details are delegated to the online appendix.

I.A. High-frequency bond data and zero curve tick-by-tick

We obtain 19 years' worth of high-frequency price data of US Treasury securities from January 1992 through December 2010 by splicing historical observations from two inter-dealer broker platforms: GovPX (1992:01–2000:12)[6] and BrokerTec (2001:01–2010:12). The merged data set covers about 60% of transactions in the secondary US Treasury bond market (Mizrach and Neely, 2006). In total, we work with around 50 million on-the-run Treasury bond quotes/transactions.

GovPX comprises Treasury bills and bonds with 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 on-the-run securities. Roughly 95% of trading occurs between 7:30AM and 5:00PM EST (see also Fleming, 1997), which we treat as the trading day. We sample bond prices at ten-minute intervals taking the last available price for each sampling point.

The raw data set contains coupon bonds. We use the equally-spaced high-frequency price data to construct the zero coupon yield curve for every sampling point following the procedure of Fisher, Nychka, and Zervos (1994).

While the availability of the high-frequency data is a restriction on the length of our sample, the 1992–2010 sample captures a homogenous interest rate environment. There is empirical

[6]GovPX started operating in 1991.

evidence that the conduct of monetary policy changed significantly during the eighties (e.g., Ang, Boivin, Dong, and Loo-Kung, 2010). The Treasury market functioning has also shifted dramatically with the advent of automated trading, interest rate derivatives and the swap market. Finally, as we discuss below, the rise of the repo market in the last two decades has had a significant influence on interest rate volatility.

I.B. Realized yield covariances

The high-frequency zero curve allows us to estimate the realized covariance matrix of yields. We consider two, three, five, seven and ten-year yields, i.e. the most liquid maturities in the secondary bond market (see also Fleming and Mizrach, 2009, Table 1).

Let y_t be the vector of zero yields with different maturities observed at time t. Time is measured in daily units. The realized covariance matrix is constructed by summing up the outer products of a vector of ten-minute yield changes, and aggregating them over the interval of one day $[t, t+1]$:

$$RCov(t, t+1; N) = \sum_{i=1,\ldots,N} \left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}} \right) \left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}} \right)'. \qquad (3.1)$$

$N = 58$ 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. For frequent sampling, the quantity (3.25) converges to the underlying quadratic covariation of yields (Jacod, 1994; Barndorff-Nielsen and Shephard, 2004). 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. In the online appendix we verify the robustness of this estimator, and compare it to alternatives proposed in the literature (e.g., Hayashi and Yoshida, 2005).

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 the $RCov$ dynamics relies on the within-day observations, excluding the volatility patterns outside the US trading hours.[7] The second choice lies in focusing on

[7]We observe several abrupt spikes in the between-day volatility, which we cannot relate to any major news in 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

intermediate and long maturities. The very short end of the curve (T-bills) is deliberately excluded from the realized covariance matrix computations because over our sample period this segment exhibited a continuing decline in trading activity, with data available only till March 2001.[8] Moreover, its dynamics is confounded by interactions with the LIBOR market and monetary policy operations. Such distortions are not directly relevant to the analysis we perform.

There are several reasons for using intraday as opposed to daily data to construct the covariance matrix of yields. First, the intraday data contain contemporaneous information about the uncertainty and the variation in liquidity demand that is hard to spot in daily returns. For instance, during the week of Lehman Brothers collapse, the realized volatility of the two-year Treasury bond based on intraday returns was almost twice as high as the daily realized volatility. Second, by the standard statistical argument the use of intraday returns increases the precision of the realized covariance matrix estimator (e.g. Andersen and Benzoni, 2008).

I.C. Empirical facts about interest rate volatility

Summary statistics. Table I reports summary statistics for weekly yields (panel *a*) and realized volatilities (panel *b*). Figure 3.1 plots average curves, both unconditional and conditional on 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, the monetary easing raises the slope of the yield curve, lifts the level of yield volatility and increases the magnitude of the hump.

[Figure 3.1 and Table I here.]

Compared to the smooth evolution of yields, the volatility curve experiences periods of elevated and abruptly changing dynamics apparent in Figure 3.2.

[Figure 3.2 here.]

---

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.

[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.

Factors in volatilities. Similar to the cross-section of yields, at least three factors are also needed to explain the dynamics of the realized volatility curve. The first three principal components explain 89.7%, 7.7% and 1.5% of its variation (panel *b* of Figure 3.3). Importantly, this observation comes from analyzing two- to ten-year yields only, and thus is not driven by idiosyncratic volatility at the very short end of the curve.

The recent financial crisis emphasizes the multivariate structure of yield volatilities. As seen in Panel *b* of Figure 3.2, while the two-year realized volatility increased visibly already during the summer 2007, the ten-year volatility remained relatively low until the Lehman collapse. Following the extraordinary measures undertaken by the Fed and the US Treasury and the promise of low rates for an "extended period of time," the two-year realized volatility fell sharply in the beginning of 2009, and reached all-time low in the second half of 2010. At the same time, as the uncertainty about the long-term economic outlook continued, the ten-year volatility remained at elevated levels.

[Figure 3.3 here.]

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, imply that the same subset of factors determines both yields and their volatilities. As one example, a single-factor CIR model suggests that the volatility is high whenever the short rate is high (Cox, Ingersoll, and Ross, 1985).

While the CIR-type prediction remained valid through the early 1980s (Chan, Karolyi, Longstaff, and Sanders, 1992), more recently the unspanned stochastic volatility (USV) literature has argued that the yield-volatility relation is weak (Collin-Dufresne and Goldstein, 2002). Using realized volatility we show that this relationship can in fact be nonlinear. Figure 3.4 scatter-plots weekly realized volatilities against the level of interest rates with a matching maturity. The shape of a nonparametric regression fitted to the data suggests an asymmetrically U-shaped relation, which contrasts with the early 1980s' episode. The volatility is low for the intermediate interest rates range, and increases when rates move toward either extreme. Naturally, after the short rate strikes the zero lower bound during the recent crisis, the volatility at the short end of the curve also dies out. Outside this special interest rate environment, the rise in volatility is more pronounced in low interest rate regimes than it is in high interest rate regimes, which explains the negative unconditional correlations between yields and volatilities reported in panel c of Table I. As an illustration, the last panel of Figure 3.4 plots the realized volatility conditional on Federal funds rate changes. We superimpose the results for the entire sample period until December 2010, and for the pre-crisis period ending in December 2007. The figure shows that the cuts in the federal funds rate lead to a stronger upward revision in volatility than do tightenings. The asymmetry is most pronounced for shorter maturities (two years) and weakens at the longer end of the curve.

[Figure 3.4 here.]

II. The model

We are interested in decomposing the variation in the yield covariance matrix in a way that is consistent with a small number of factors driving yields. To this end, we formulate a flexible model able to describe the joint dynamics of yields and volatilities. Later, we verify its econometric fit and provide an economic interpretation of factors. All proofs are collected in the online appendix.

Our benchmark model is cast in a reduced-form continuous-time framework. We distinguish between two types of factors: expectations factors X_t, and covariance factors V_t. The physical dynamics are given by the system:

$$dX_t = (\mu_X + \mathcal{K}_X X_t)dt + \sqrt{V_t}dZ^{\mathbb{P}}_{X,t} \quad (3.2)$$

$$dV_t = (\Omega\Omega' + MV_t + V_t M')\,dt + \sqrt{V_t}dW^{\mathbb{P}}_t Q + Q' dW^{\mathbb{P}\prime}_t \sqrt{V_t}, \quad (3.3)$$

where X_t is a n-vector, and V_t is a $n \times n$ a covariance matrix process (Bru, 1991; Gourieroux, Jasiak, and Sufana, 2009a). Accordingly, $Z^{\mathbb{P}}_X$ and $W^{\mathbb{P}}$ are an n-dimensional vector and a $n \times n$ matrix of independent Brownian motions.[9] μ_X is a n-vector of parameters and $\mathcal{K}_X$, M

[9]It is straightforward to introduce correlations between expectations and volatility shocks Z_X and W by setting $dZ_X = dW\rho + \sqrt{1-\rho'\rho}dB$ 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 online appendix. As a preliminary check, we can learn how these shocks in yields and volatilities are related in the data by estimating a VAR(1) for the joint system. Let us include three bond portfolios mimicking the level, slope and curvature of the yield curve plus three realized volatility factors: the level RV^{2Y}_t, the slope $(RV^{10Y}_t - RV^{2Y}_t)$ and the covariance $RCov^{5Y,10Y}_t$. The VAR results are not reported in the main text. The highest cross-correlation between shocks does not exceed 16% (the yield slope portfolio and the volatility level). Therefore, for the empirical implementation of the model, we set $\rho = 0$. The details of model estimation follow in Section II.B.

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:

$$r_t = \gamma_0 + \gamma_X' X_t + \gamma_f f_t. \tag{3.4}$$

The state f_t evolves as:

$$df_t = (\mu_f + \mathcal{K}_f f_t + \mathcal{K}_{fX} X_t)dt + \sigma_f dZ^{\mathbb{P}}_{f,t}, \tag{3.5}$$

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.

One can think of f_t as a short-term monetary policy expectations factor. X_t, instead, summarizes economic variables that drive longer term yields. As such, X_t can impact the conditional expectation of f_t. Time-varying volatility enters the model through long term factors: V_t describes the amount of risk in the economy, with out-of-diagonal elements of V_t determining the conditional mix between X_t's. We focus on modeling the stochastic covariance of yields of intermediate to long maturities, whose variation we can observe with support of high frequency data. Given the different institutional properties of shortest maturity yields and the unavailability of traded high-frequency data, we assume constant conditional volatility for f_t.

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:

$$dY_t = (\mu_Y + \mathcal{K}_Y Y_t)\, dt + \Sigma_Y(V_t) dZ^{\mathbb{P}}_t, \tag{3.6}$$

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 are priced using the standard no-arbitrage argument. We specify a general reduced-form compensation $\Lambda_{Y,t}$ required by investors to face shocks in the state vector:

$$\Lambda_{Y,t} = \Sigma_Y^{-1}(V_t)\left(\lambda_Y^0 + \lambda_Y^1 Y_t\right), \tag{3.7}$$

where λ_Y^0 is a $(n+1)$-vector and λ_Y^1 is a $(n+1)\times(n+1)$ matrix of parameters. We require the invertibility of each block of matrix $\Sigma_Y(V_t)$, which is ensured by $\sqrt{V_t}$ positive definite, and $\sigma_f > 0$. In equation (3.7), we assume that only Z shocks are priced. Therefore, the risk neutral dynamics of X_t follow from the standard drift adjustment:

$$\mu_Y^{\mathbb{Q}} = \mu_Y - \lambda_Y^0 \tag{3.8}$$

$$\mathcal{K}_Y^{\mathbb{Q}} = \mathcal{K}_Y - \lambda_Y^1, \tag{3.9}$$

and the dynamics of V_t remain unchanged.[10]

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

$$P(t,\tau) = e^{A(\tau)+B(\tau)'Y_t + Tr[C(\tau)V_t]}, \tag{3.10}$$

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

$$\frac{\partial A(\tau)}{\partial \tau} = B(\tau)'\mu_Y^{\mathbb{Q}} + \frac{1}{2}B_f^2(\tau)\sigma_f^2 + kTr\left[Q'QC(\tau)\right] - \gamma_0 \tag{3.11}$$

$$\frac{\partial B(\tau)}{\partial \tau} = K_Y^{\mathbb{Q}\prime}B(\tau) - \gamma_Y \tag{3.12}$$

$$\frac{\partial C(\tau)}{\partial \tau} = \frac{1}{2}B_X(\tau)B_X(\tau)' + C(\tau)M + M'C(\tau) + 2C(\tau)Q'QC(\tau), \tag{3.13}$$

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 (3.11)–(3.13) are $A(0) = 0_{1\times 1}, B(0) = 0_{(n+1)\times 1}$ and $C(0) = 0_{n\times n}$. $B(\tau)$ has a standard solution typical to Gaussian models; $C(\tau)$ 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:

---

[10] It is straightforward to introduce a priced volatility risk without losing the tractability. However, due to weak spanning of volatility states by bonds it is 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, which is not the focus of this paper.

$$y_t^\tau = -\frac{A(\tau)}{\tau} - \frac{B(\tau)'}{\tau}Y_t - Tr\left[\frac{C(\tau)}{\tau}V_t\right]. \tag{3.14}$$

Yields are an affine function of the entire state vector $(Y_t, \mathrm{vec}(V_t)')'$. Under uncorrelated shocks dZ and dW, and the short rate (3.4), 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 (3.6). Thus, the instantaneous yield covariation is only driven by the covariance factors:

$$\begin{aligned} v_t^{\tau_i,\tau_j} &:= \frac{1}{dt}\langle dy_t^{\tau_1}, dy_t^{\tau_2}\rangle \\ &= \frac{1}{\tau_1\tau_2}\{Tr\left[B_X(\tau_2)B_X(\tau_1)' + 4C(\tau_2)Q'QC(\tau_1)\right]V_t + B_f(\tau_1)B_f(\tau_2)\sigma_f^2\}. \end{aligned} \tag{3.15}$$

II.A. Discussion

In the basic setup, we consider three variables in Y_t, i.e. f_t plus a two-dimensional vector X_t, where X_t has a 2×2 covariance matrix V_t. The form of V_t leads to a three-factor model of yield volatilities with a covariance and two volatility states. The combination of six factors gives us the flexibility to fit both yields and their volatilities. The state space we consider is large, but involves a manageable number of identified parameters (13 excluding $\Lambda_{Y,t}$).

Our distinction between volatility and yield curve variables builds on the $A_m(n)$ classification of ATSMs introduced by Dai and Singleton (2000). The new element is that V_t represents a complete covariance matrix, and as such involves an interaction component which can switch sign.[11] To make the roles of factors interpretable, we exclude interactions between V_t and X_t through the drift. Even though such interactions are usually allowed, we find that they are not called for by the data.

The presence of V_t in expression (3.14) distinguishes our model from the unspanned stochastic volatility settings, which impose explicit restriction so that the volatility factors do not enter the cross section of yields. Such separation usually improves the volatility fit of low-dimensional ATSMs (Collin-Dufresne, Goldstein, and Jones, 2009). However, as highlighted by the literature, there appear to be few reasons, except statistical ones, for such constraint to strictly hold. The volatility variables could appear in the term structure of yields through at least two channels. One of them is the convexity bias (see also Phoa, 1997; Joslin, 2010). A

[11] In ATSMs, independent CIR processes generate stochastic volatility one-by-one, and therefore conditional covariances they imply are a linear combination of the volatility factors.

second one is the relation between the amount of uncertainty and the expected excess returns. In that the term premia compensate for risks, they should be related to the changing amount of interest rate volatility.

II.B. Estimation

We estimate the model on a weekly frequency ($\Delta t = \frac{1}{52}$) combining pseudo-maximum likelihood with filtering. All technical details on estimation and identification are presented in the online appendix.

The transition equation for Y_t is specified as an Euler approximation of the physical dynamics (3.6):[12]

$$Y_{t+\Delta t} = \bar{\mu}_{Y,\Delta t} + \Phi_{Y,\Delta t} Y_t + u^Y_{t+\Delta t}, \tag{3.16}$$

where u^Y_t is a vector of heteroskedastic innovations $u^Y_t = \Sigma_Y(V_t)\sqrt{\Delta t}\epsilon_{t+\Delta t}$. The transition equation for the matrix process V_t is obtained by an exact discretization of the dynamics (3.3):

$$V_{t+\Delta t} = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} + u^V_{t+\Delta t}, \tag{3.17}$$

where u^V_t represents a symmetric matrix of heteroskedastic innovation. $\bar{\mu}_{Y,\Delta t}$, $\mu_{V,\Delta t}$, $\Phi_{Y,\Delta t}$ and $\Phi_{V,\Delta t}$ are explicit functions of the continuous-time parameters.

We introduce two types of measurement equations based on yields (y^τ_t) and their quadratic covariation (v^{τ_i,τ_j}_t):

$$y^\tau_t = f(S_t;\Theta) + \sqrt{R_y} e^y_t \tag{3.18}$$
$$v^{\tau_i,\tau_j}_t = g(V_t;\Theta) + \sqrt{R_v} e^v_t. \tag{3.19}$$

Functions $f(S_t;\Theta)$ and $g(V_t;\Theta)$ denote model-implied expressions (3.14) and (3.15) corresponding to the observed measurements y^τ_t and v^{τ_i,τ_j}_t; Θ collects model parameters. v^{τ_i,τ_j}_t is obtained from the high-frequency zero curve using estimator (3.25).

[12] In the online appendix, 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 (3.16).

The weekly realized covariance matrix estimator is still noisy. Therefore, 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 with $R_y = \sigma_y^2 I_6$. Additionally, we include three volatility measurements: variances of the two- and ten-year bond and the covariance between the five- and ten-year bond. We allow for different errors across equations, i.e. $R_v = \text{diag}(\sigma_v^i)$, where $i = 1, 2, 3$, and $\text{diag}(\cdot)$ is a diagonal matrix.

The assumption that the realized and the instantaneous covariance matrices of yields are equivalent is an approximation implicit in our choice of measurements. Equation (3.19) uses the fact 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 absence of jumps, the estimator (3.25) converges to the integrated covariance matrix of yields. We recognize that the measurement (3.19) is not exact, but on a weekly frequency the error due to the approximation can be assumed negligible.

In order to handle the non-Gaussianity in factor dynamics we use the square-root Unscented Kalman Filter (UKF), proposed by Julier and Uhlmann (1997), and recently applied by e.g. Carr and Wu (2007) or Christoffersen, Jacobs, Karoui, and Mimouni (2009). To secure against local minima, we apply a global optimization algorithm—differential evolution (Price, Storn, and Lampinen, 2005). We confirm that the algorithm achieves the global minimum by repeating the estimation several times.

Finally, we impose parameter restrictions that ensure econometric identification of the model. After identification restrictions, our preferred model that we discuss below 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.

With the inclusion of several parameters in $\Lambda_{Y,t}$, the search for a global maximum likelihood optimum becomes slow. For this reason, we adopt a parsimonious form of market prices of risk by estimating only two diagonal parameters in $\Lambda_{Y,t}$ related to longer-duration factors X_t, $(\lambda_{11}^{1X}, \lambda_{22}^{1X})$, i.e. those that carry information about bond risk premium. The goal of our model is to decompose the cross section of yields and volatilities in a mutually consistent way. Estimating alternative and richer forms of the market price of risk (unreported), we find that the risk premium specification has a minor impact on the factor decomposition we obtain. This observation is consistent with Collin-Dufresne, Goldstein, and Jones (2009).

III. Estimation results

III.A. Parameter estimates

Table II provides parameter estimates for our model indicated by the label G_3SV_3 (meaning: three conditionally Gaussian plus three volatility factors). On the level of explaining yields, we benchmark the performance of the estimated model to the Gaussian three-factor counterpart (G_3SV_0) which we estimate with the standard Kalman filter. This comparison evaluates whether volatility factors, while beneficial for capturing second moments of yields, introduce a risk of misspecification on the side of fitting yields.

[Table II here.]

For each factor type (V_t and Y_t) at least one variable is more persistent and one faster-moving. While this observation is common for the yield dynamics, our estimates indicate that it also holds true for the volatility curve. The most important difference between V_t states comes in their respective volatilities (vol of vol), as captured by the parameters of the Q matrix. The volatility of the V_{11} and V_{12} state is just about a half of V_{22}, indicating that a multivariate volatility assumption is borne out 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. The degrees of freedom parameter k in equation (3.3) controls the extent of non-Gaussianity present in the volatility factors: the higher the k the more Gaussian the dynamics.[13] Across various version of the model which we estimated, the model consistently selects $k = 2$. This low value reflects the need for factors with a pronounced right tail.

The bottom section of Table II provides the log-likelihood values for estimated specifications. For comparison with the Gaussian benchmark, we split the log-likelihood implied by our

[13] Buraschi, Cieslak, and Trojani (2008) discuss the role the degrees of freedom parameter in the V_t dynamics. See also online appendix 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.

model into two parts related to yields and to volatilities, respectively. The numbers indicate the improvement in volatility modeling offered by G_3SV_3 does not impede the fitting of interest rates across maturities.

Simultaneously, the model is able to explain well the evolution of the second moments. Figure 3.5 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 volatility very closely in terms of magnitudes, persistence and signs, both at the short and the long end of the curve.

[Figure 3.5 here.]

Table III confirms these conclusions. We provide two measures 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 three basis points. Its performance is very similar to the purely Gaussian case. At the same time, we are able to explain 98% of the variation in the second moments dynamics. An important result is that the model captures volatilities equally well at long and short maturities (Table IV), explaining 99% of two-year volatility and 97% of the ten-year volatility. Recently several authors have noticed that fitting yield volatilities at the long maturity range remains a significant challenge; even at short maturities the fit can be poor with traditional models explaining only about 50% of the volatility variation. Our results show that a multifactor volatility combined with the information in the realized covariance matrix of yields successfully alleviates these problems.

[Table III here.]

One might worry about over-fitting in our six factor model either on the yield or on the volatility side. To address this concern, we estimate the G_3SV_3 model on the sample from January 1992 through December 2007, and then use these parameter to price yields and fit volatilities on the whole sample including the financial crisis, i.e. January 2008 through

December 2010. These results are reported in the last column of Table II. It turns out that the model can fit both yields and volatilities in that period equally well as in the pre-crisis period. This result is reassuring especially because the parameters are stable during the financial crisis which saw the most extreme movements in volatility on record.

III.B. Filtered states

This section discusses the properties and roles of the factors filtered by the model. Their good fit to the data indicates that these six factors summarize well the information in the cross-section of yields and volatilities.

Figure 3.6 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 volatility factors $\bar{V} = (V_{11}, V_{12}, V_{22})'$.

[Figure 3.6 here.]

Factors have a joint interpretation: The roles that the model assigns to the X_t states correspond 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. Figure 3.7 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 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. 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 Federal Funds rate (see Figure 3.6c). 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.[14]

[Figure 3.7 here.]

Corresponding with the X_1 and X_2 factors, V_{11} generates volatility at longer maturities, while V_{22} captures the shorter end of the volatility curve. Table IV 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 across maturities has an opposite pattern: The effect of V_{22} decreases whereas the effect V_{11} increases with the maturity. The out-of-diagonal element V_{12} is strongly significant, and has the largest impact on the yield covariance dynamics. Accordingly, we will term the volatility factors as the long-end (long-duration) volatility state, the short-end (short-duration) volatility state and a covariance state, respectively.

[Table IV here.]

III.C. Are volatility factors revealed by the yield curve?

Let us focus again on Figure 3.7. Its right-hand panels $(b1, b2, b3)$ plot the response of the yield curve to large changes in volatility. 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—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 is intuitive: 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 of the yield curve to V_{22} should be less pronounced, or even negligible.

[14]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 federal funds rate.

III.D. Zero rate bound

How may our factor interpretation change in a zero-lower bound environment that the US entered post 2008? In the last two years of our sample the effective federal funds rate (FFR) has been fluctuating between zero and 25 basis points. At the longer end, however, yields have been volatile and away from the zero bound. Our model accommodates such a yield environment due the maturity-ordered factor structure in yields and the stochastic covariance matrix of yield curve shocks. Throughout the sample, the factor f_t tracks short-term monetary policy expectations and is very smooth, which can be attributed to the transparency of Fed's communication. The estimated volatility of shocks to factor f_t, given by σ_f in Table II, is identical pre-crisis (1992-2007) and in the full sample (1992-2010). Thus, there is no change in statistical and economic role of f_t before and during the period when the zero lower bound was hit. The long-end factor X_1 has a zero loading at short maturities and thus is little affected by the zero bound. Over the course of recent crisis, the Fed has been increasingly clear at communicating that FFR target range will stay at very low levels for an extended period. This expectations channel has brought the short term volatility V_{22} to historically lowest levels following the Lehman failure, and again in the second half of 2010 when the Fed started pursuing the second round of quantitative easing (QE2). At the same time, on both occasions long-end volatility V_{11} has ticked up (see Figure 3.8).

IV. Which economic forces are reflected in the Treasury volatility?

This section explores the link between the volatility factors and various proxies for liquidity, macroeconomic uncertainty and alternative measures of market volatility. We also ask how yield volatility states relate to term premia in Treasury bonds. We show that a model-based decomposition of Treasury volatility facilitates our understanding of its economic content.

There are several takeaways from this section. First, separating the volatility of Treasuries into the short- and long-duration factors is essential for revealing its economic content. Second, the volatility factors provide timely information about the state of the economy that cannot be read from yields at low frequencies. Third, a large part of the variation in volatility can be traced back to time-varying liquidity conditions and macroeconomic uncertainty. Finally, interest rate volatility provides supportive evidence on the presence of hidden factors driving bond risk premia, and suggests that the hidden factors can reflect the uncertainty about monetary policy.

IV.A. Short versus long-end volatility dynamics

Figure 3.8 displays the evolution of long- and short-end volatility in our sample, showing that volatility factors have a pattern that is distinct from the business cycle variation. While typically before recessions only the short-end volatility rises, after recessions short- and long-end volatility both tend to spike up. The latter is the case for each post-recession period in our sample: in 1992 (post 1990–1991 recession) the long-end volatility factor increased by more than six standard deviations away from its unconditional mean; similarly in 2003 (post 2001 recession) long-end volatility moved six standard deviations from the mean; at the end of 2008 (post 2007–2008 recession) the short-end volatility factor V_{22} reached all time high moving more than eight standard deviations from its unconditional average level. For comparison, the unconditional mean of the volatility is around 90 basis points. The short-term volatility factor V_{22} also increases sharply during periods of distress in asset markets such as the LTCM and the Russian crisis or the dot-com bubble. Interestingly, during the recent financial crisis, the short-term volatility reached extreme levels already in the Fall 2007 and persisted until 2009; the long-term volatility, instead, remained low until the bankruptcy of Lehman Brothers. Possibly, prior to September 2008 market participants considered the ongoing crisis to be short lasting similar to the LTCM crisis, and the fall of Lehman changed that view.

[Figure 3.8 here.]

The advantage of the model-based decomposition over direct volatility proxies such as realized or implied volatility is that the latter merge together effects of different duration. This is evident by comparing the filtered volatility factors (Figure 3.8) with the realized volatility (Figure 3.2). Later in this section, we discuss the link between our decomposition and implied yield volatility from options.

IV.B. Links between volatility and liquidity

A store of liquidity and the highest quality of collateral are the main motives driving demand for the US Treasuries. We ask how these motives are reflected in yield volatility by studying three types of linkages between liquidity and volatility.

The first is the financial conditions and funding liquidity channel. High value of the long-end volatility factor V_{11} is observed when prospects for the economy are weak, and monetary policy is expected to ease (see Section IV.C). In such periods, financial conditions and funding liquidity typically deteriorate. Hence, one would expect the long volatility state V_{11} and funding liquidity to be closely linked. To reconcile this interpretation with the data, we associate the volatility factor V_{11} with a measure of funding liquidity conditions extracted from the cross-section of on- and off-the-run bonds, as recently proposed in Fontaine and Garcia (2011):[15]

$$\mathrm{liq}_t^{FG} = \delta_0 + \delta_1 V_{11,t} + \varepsilon_t. \tag{3.20}$$

The regression of the funding liquidity factor liq^{FG} on the long volatility factor V_{11} gives an R^2 of 39% with V_{11} being negatively related to the value of funding liquidity and highly significant (Newey-West t-statistic is -5.2). Elevated levels of long-term volatility signal a decrease in the value of funding liquidity (e.g. as induced by monetary policy easing). However, to uncover this strong relationship one needs the model-based decomposition of volatility: A regression of liq^{FG} on the raw realized volatility or the short-end volatility factor V_{22} gives R^2's of about 10–11% and an insignificant t-statistic.

It is interesting to study the lead-lag relationship between the long-end volatility and the funding liquidity. To investigate the lead-lag structure, we modify the regression (3.20) to include lags:

$$\mathrm{liq}_t^{FG} = \delta_{0,i} + \delta_{1,i} V_{11,t-i} + \varepsilon_t, \tag{3.21}$$

with $i = -18, \ldots, 18$ months. Panel a of Figure 3.9 shows the R^2's, $\delta_{1,i}$ coefficient and the t-statistics from these regressions. The results suggest a strong predictive power of the long-end volatility for funding liquidity up to several months ahead. An increase in long volatility state signals a lower value of funding liquidity with an R^2 close to 50%. Thus, the long-end volatility factor can be viewed as a leading indicator for funding liquidity conditions.

[15] Fontaine and Garcia (2011) filter the funding liquidity factor using on-the-run and and off-the run coupon bond yields. Specifically, their factor measures the value of funding liquidity, i.e. an increase in the liquidity factor means an increase in the value of funding liquidity and thus tighter funding liquidity conditions. We thank Jean-Sebastien Fontaine for sharing the data with us.

[Figure 3.9 here.]

Second, the level of liquidity can be associated with the volatility through the (slow-moving) arbitrage capital channel. If the arbitrage capital is abundant, it holds prices of Treasuries tightly aligned even intraday, keeping their volatility low. The relationship between the availability of arbitrage capital and liquidity in Treasury market has been studied by Hu, Pan, and Wang (2013). Their noise illiquidity factor, which measures temporary price deviations from the efficient yield represented by the Nelson-Siegel-Svensson model, is shown to track the fluctuations in overall market liquidity.[16] We relate the noise illiquidity measure to our three volatility factors. Panel *a* of Table V reports the weekly regression results. Volatility factors explain 25% of the variation in the noise illiquidity factor in our sample. All three factors are significant, showing that illiquidity relates to the variation in the entire covariance matrix of yields. This is intuitive as the illiquidity measure aggregates the deviations across the whole curve and does not distinguish between the short- and long-end noise. Nevertheless, the comparison of R^2's in the last row of panel *a* in Table V suggests that short-end volatility factor V_{22} contributes most to the explained variation in illiquidity measure.

[Table V here.]

Similar to equation (3.21), we run the following lead-lag regression of noise illiquidity:

$$\text{illiq}_t = \delta_{0,i} + \delta_{1,i} V_{22,t-i} + \varepsilon_t. \tag{3.22}$$

Panel *b* of Figure 3.9 plots the regression results. The results indicate a strong positive link whereby the short-end volatility factor predicts the subsequent illiquidity up to ten months ahead. High levels of the short-end volatility factor predict that the arbitrage capital will be scarcer and the Treasury market less liquid.

Third, we explore the linkages between the volatility and liquidity through the collateral channel. Large portion of outstanding Treasuries is used by investors as collateral in repurchases or derivative OTC transactions. In periods of large demand for liquid collateral,

[16]Hu, Pan, and Wang (2013) define the noise illiquidity measure as the root mean squared yield pricing error on a given day. Every day, they estimate a Nelson-Siegel-Svensson model using the CRSP Treasuries dataset. Higher value of illiquidity signals the shortage of arbitrage capital which would otherwise align the prices of individual bonds.

market participants may choose to fail to deliver on their repo transaction.[17] Because the penalty for the failure is linked to the short term interest rate, such a situation occurs especially when the interest rates are low.[18] The variation in the volumes of repo failures involving Treasury bonds as collateral should then be informative about the demand pressure in the Treasury market. Given that such demand pressures tend to be short term, we would expect the failures to be visible mainly in the short term volatility.[19] We obtain weekly data on the dollar value of failures in Treasuries from the New York Fed website and study how they relate to the volatility factors.[20] Although given the likely endogeneity between interest rate volatility and fails, it may be more natural to regress volatility on fails, we reverse the order and include the three volatility factors on the right hand side of the regression. Panel *b* of Table V summarizes the regression output. As expected, failures to deliver have the strongest link to the short volatility factor V_{22}. Three volatility factors explain 31% of the variation in failures.

In summary, the results in this section point to strong links between the Treasury volatility and liquidity. The ability of volatility factors to predict the variation in funding liquidity stresses that important information can be read directly from on-the-run yields. However, while the volatility states aggregate information only from on-the-run Treasury yield curve, this information is difficult to extract by observing yields at infrequent intervals.

IV.C. Volatility and macroeconomic uncertainty

This section asks how much of the variation in interest rate volatility is due to the changing macro uncertainty. We relate the volatility factors to survey-based proxies for expectations and uncertainties about selected macro variables.

[17]Since May 2009, when a special charge for the failure was introduced, failures to deliver decreased to extremely low levels. See Garbade, Keane, Logan, Stokes, and Wolgemuth (2010) for more details.

[18]Failure to deliver can happen for other reasons such as mis-communication between the two counterparties. However, Fleming and Garbade (2005) argue that most of the failures are related to other factors.

[19]Alternatively, we could use the variation in haircuts applied to Treasury collateral to study the collateral channel—an increase in Treasury volatility lead to an increase in haircuts applied to Treasury collateral which trigger margin calls and thus decrease the liquidity in the market. Due to the lack of data on haircuts we analyze only the failures to deliver.

[20]The New York FED publishes the dollar amounts of settlements failures for various types of bonds (Treasury, agency, corporate bonds and MBS) on weekly basis. The data are available starting from July 1990.

Our monthly survey data are spliced from two sources. To obtain forecasts of macro variables, we use the BlueChip Economic Indicators (BCEI) survey and BlueChip Financial Forecasts (BCFF). As evident in the transcripts of the FOMC meetings, these surveys are regularly used by policymakers at the Fed to read market expectations. The online appendix provides details about the survey data including their timing within a 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.

The macro variables that 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 highly informative about our filtered states. We are able to explain up to about 90% and 30% of variation in the yield curve and volatility factors, respectively.

Table VI presents the regressions of X_t, f_t, V_t factors on consensus and uncertainty proxies.[21] The loadings on the yield curve factors confirms their distinct responses to the economic environment. The long-term yield factor X_1 is driven by the expectations about the inflation, with one-year ahead inflation consensus forecast emerging as its key mover. The real part is represented by the expectations about RGDP and UNEMPL both of which help explain the variation in long term yields. The intermediate factor X_2 is related to the expectations and uncertainty about the future path of the the monetary policy. While the uncertainty about industrial production increases intermediate yields, the uncertainty about monetary policy and housing starts both decrease yields in this segment of the curve. Finally, in addition to being almost completely explained by the expectations about monetary policy, the short-end factor f_t responds somewhat to the expected CPI inflation and the current stance of the real activity (expectations about IP). Not surprisingly, monetary policy is an important driver of the short and intermediate yields (factors X_2 and f_t) but its impact diminishes with the state's maturity, telling us how the monetary impulse is transmitted to the longer segments of the curve. Monetary policy variable is not significant in the X_1 regression; the long end of the curve is almost entirely determined by the macroeconomic rather than monetary forces.

[21] To obtain the results presented in Table VI, we select the regressors based on their significance.

[Table VI here.]

Volatility states reflect predominantly the uncertainty about the macroeconomic outlook. Each of the two volatility factors V_{11} and V_{22} captures a distinct domain of uncertainty. Indeed, variables inducing long- versus short-run volatility movements form essentially separate sets. The long-end volatility V_{11} is pronouncedly influenced by the uncertainty about unemployment and housing starts. The short-end volatility V_{22} is most strongly linked to the uncertainty about monetary policy (t-statistic of 4.37), and to a weaker but still significant degree, to the uncertainty about inflation and industrial production. Notably, shocks to these variables are almost contemporaneously observed and rather short-lived, thus providing a contrast to more sticky metrics such as the real GDP.[22]

[Figure 3.10 here.]

To better understand the difference between the volatility states, in panels a through e of Figure 3.10 we plot their impulse-responses to various sources of macro uncertainty. The results are obtained on the pre-crisis sample 1992-2007 to avoid a large spike in volatility at the time of the Lehman bankruptcy which could potentially drive the results. The respective volatility responses 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 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.

[22]Ranking inflation uncertainty as short-lived may be surprising. This can be partially explained by the sample period we use (1992:01–2010:12). 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.

We find that under our identification scheme,[23] the covariance state V_{12} is positively linked to the realized correlation of the long-to-medium part of the yield curve (see Figure 3.2, panel *c*). Conditional on this observation, we can interpret its signs in survey regressions. V_{12} is associated with uncertainty measures about the unemployment, inflation, and monetary policy. An increase in uncertainty about the unemployment and monetary policy implies a decline in V_{12}, and a lower correlation between the long and medium part of the yield curve. V_{12} is also related to the expectations about the real GDP: An increase in growth expectations lowers the correlations along the curve. Intuitively, the different segments of the yield curve 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 one of the two NBER-proclaimed recessions in our sample. V_{12} captures precisely those moves.

IV.D. Are volatility factors driving bond risk premia?

A growing literature explores the variation in bond risk premia focussing on so-called "hidden" factors that have a negligible effect on the cross section of yields but play an important role in forecasting bond returns (Duffee, 2011). As we document above, the volatility factors lend themselves as natural candidates to the interpretation as hidden states because they are imperceptible from the cross section of yields. An important question is whether and how yield volatility is reflected in bond risk premia, and whether it contributes to the bond return predictability beyond other strong return forecasters. We run the following predictive regression:

$$rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 \widehat{cf}_t + \gamma_2 V_{11t} + \gamma_3 V_{22t} + \varepsilon_{t+1}^{(n)}, \tag{3.23}$$

where $rx_{t+1}^{(n)}$ denotes the one-year bond excess return for a bond with a n-year maturity and $\widehat{cf}$ denotes the return-forecasting factor proposed in Cieslak and Povala (2011). $\widehat{cf}$ captures the term premium that is most pronounced in the variation of long-term bonds.

[Table VII here.]

[23] 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.

Table VII reports the results of estimating equation (3.23). We also include separate results using only $\widehat{cf}$ and only V_{11} and V_{22} as return predictors. In the presence of the return-forecasting factor, only the short-end volatility V_{22} adds independent explanatory power about the variation in bond excess returns, and mostly so at short maturities. Its contribution to the regression R^2 is 6% for the excess return on a two-year bond with a significant t-statistic of 2.84. Recall that it is the uncertainty about the near-term monetary policy what drives the largest portion of the short-end volatility (last column of Table VI). In line with our intuition, the volatility factor V_{22} loads positively on bond excess returns, meaning that investors require higher returns when the monetary policy uncertainty and thus the short-end volatility is high. One standard deviation increase in the short-end volatility is associated with 0.3 standard deviations increase in two-year bonds excess return. This link between volatility and bond risk premium advances our understanding of the nature and sources of hidden factors in term premia.

IV.E. Volatility factors versus implied volatility measures

An established measure of volatility in Treasury market is the Merrill Option Volatility Estimate (MOVE) index which is a yield curve weighted index of the normalized implied volatility on one-month Treasury options.

A common feature of implied volatility measures such as CBOE VIX or MOVE is that they are constructed from options on securities with a stream of cash flows. Hence, it is difficult to separate the sources of volatility using these series. MOVE, for instance, is the weighted average of implied volatilities from options on the Treasury coupon bonds with maturities of two, five, ten and 30 years.[24]

To understand how MOVE relates to different factors in volatility, we run the following regression at a weekly frequency:

$$\text{MOVE}_t = \alpha_0 + \alpha_1 V_{11,t} + \alpha_2 V_{12,t} + \alpha_3 V_{22,t} + \varepsilon_t. \tag{3.24}$$

Table VIII summarizes the results, and Figure 3.11 plots the MOVE index against the regression fit.

[24] The maturities are weighted as follows: 20% two-year, 20% five-year, 40% ten-year and 20% 30-year bond. These allocations are based on estimates of option trading volumes in each maturity segment.

[Table VIII and Figure 3.11 here.]

With an R^2 of 68%, three volatility factors explain a large part of the weekly variation in MOVE. For comparison, two-year realized volatility and the VIX index explain 47% and 36% of the variation in MOVE, respectively. To assess the relationship with the equity implied volatility, we add the VIX index to the regression (last column in panel *c* of Table VIII). All of the volatility factors contribute significantly to the explained variation in MOVE and remain equally significant in the presence of VIX. Especially, the contribution of the covolatility factor V_{21} to the explained variation in MOVE is substantial. The results of regression (3.24) suggest that the realized and implied volatilities track each other closely with the deviations being short-lived and small in magnitude. There is only one period of prolonged divergence in 2006-2007—a period marked with extremely low volatility and high liquidity.

V. Conclusions

We explain the drivers of the riskiness of the Treasury yield curve over time and across maturities. We combine 19 years' worth of tick-by-tick bond transaction data with a term structure model to identify two distinct components of volatility and a factor capturing the comovement of yields. The short-duration volatility is pronouncedly related to the uncertainty about the monetary policy and demand for liquid collateral. The long-duration volatility reflects the uncertainty about the real activity and funding liquidity conditions. Throughout, we stress the importance of the model-based decomposition of Treasury volatility for studying its economic content. Its usefulness becomes particularly evident against the backdrop of the low interest rate environment in which the different maturity segments of the yield curve tend to move with an increasing autonomy.

VI. Figures

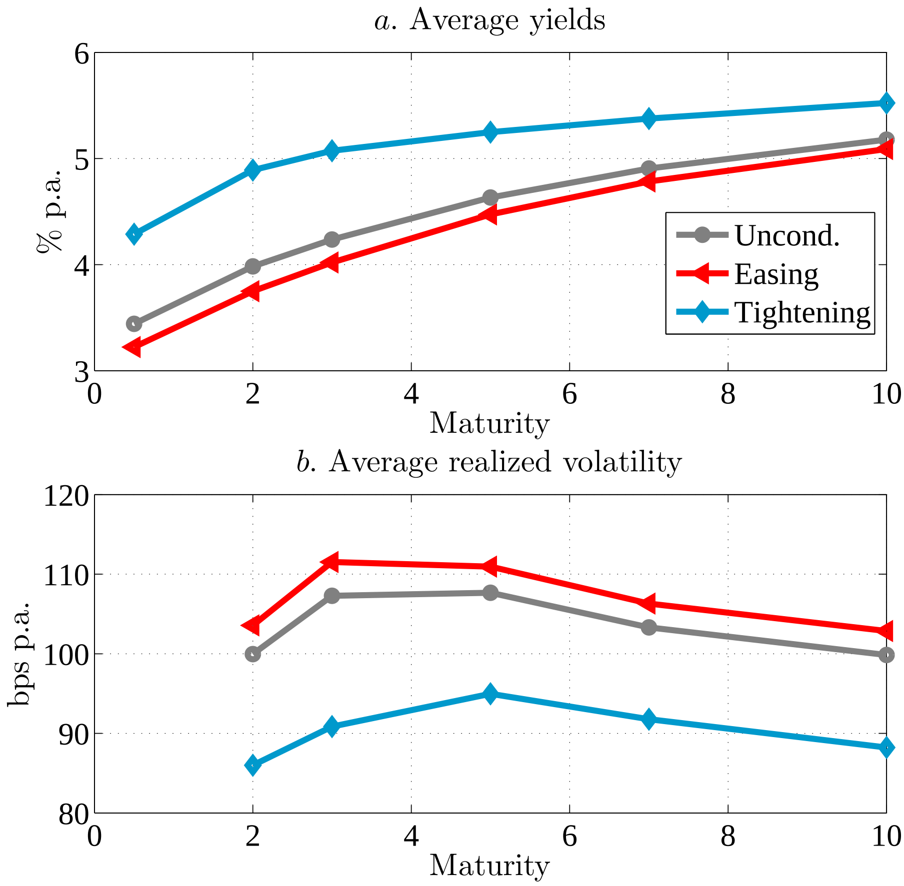


Figure 3.1: Yield and volatility curves

The figure plots the yield and volatility curves: unconditional mean for the whole sample 1992:01–2010:12, 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. The realized volatility curves are computed from the weekly data and annualized ($\times 52$).

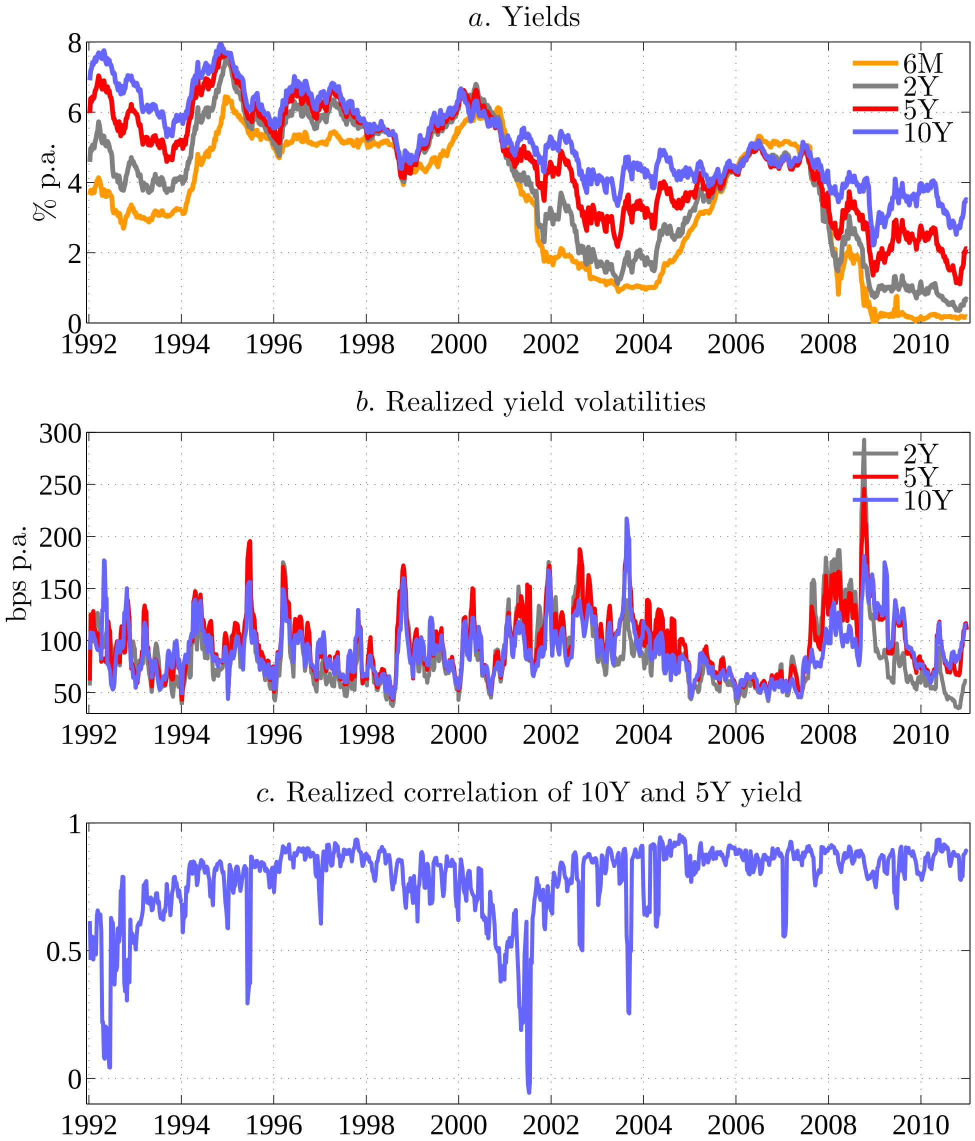


Figure 3.2: Evolution of yields and volatilities

The figure plots the dynamics of weekly yields, realized volatilities and correlation over the 1992:01–2010:12 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.

VII. Tables

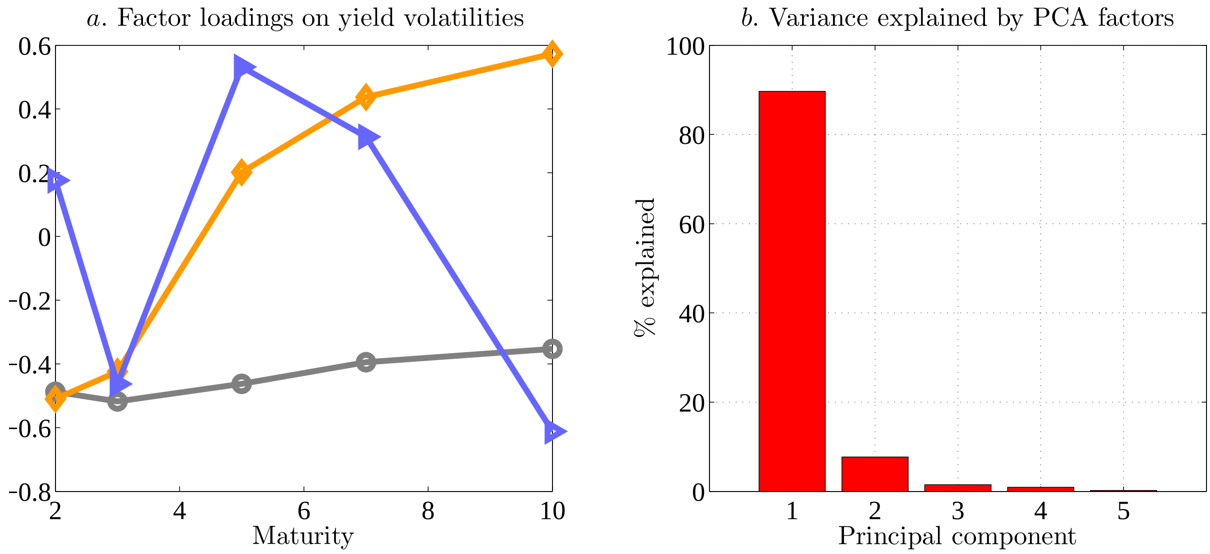


Figure 3.3: Factors in yield volatilities

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.

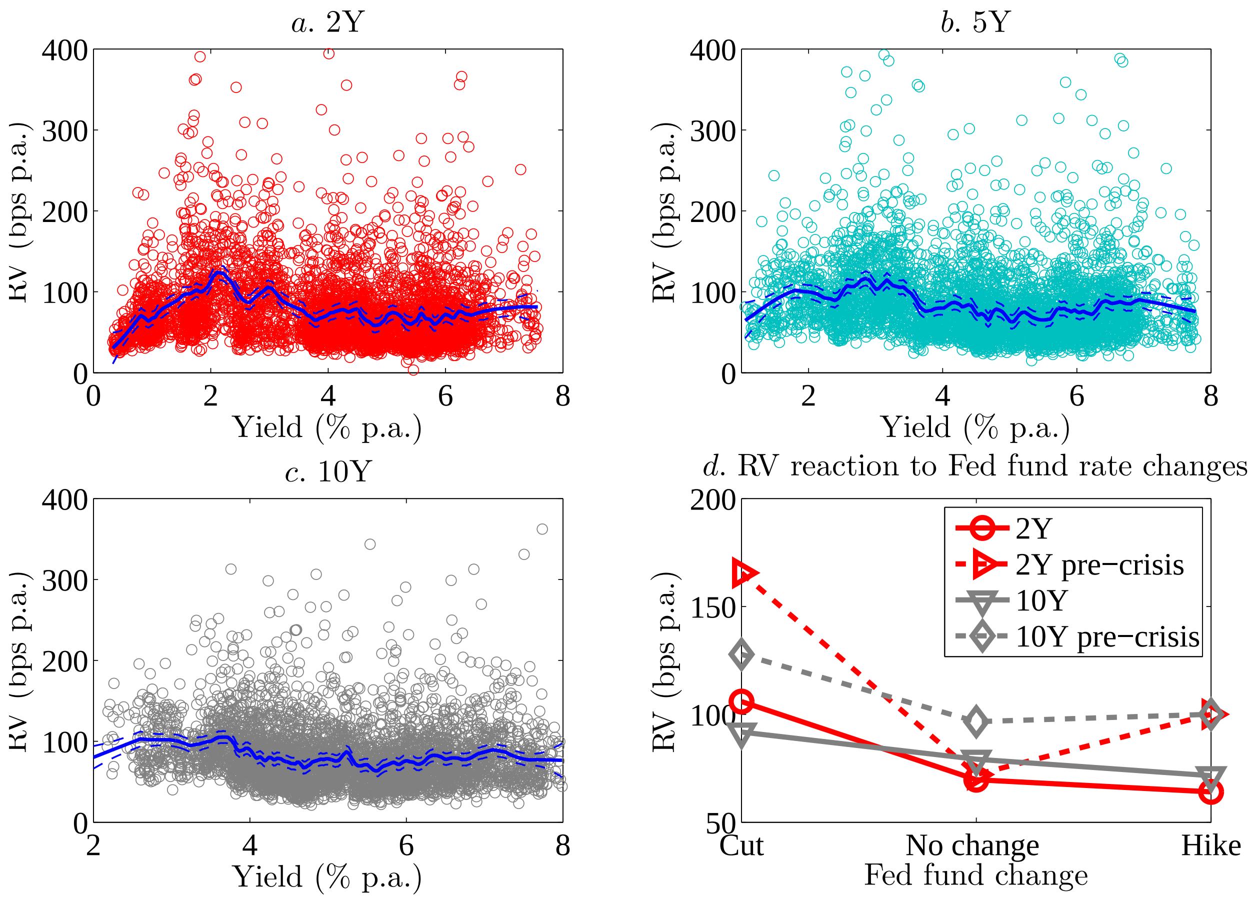


Figure 3.4: Yield level versus yield volatility

The figure depicts the relationship between yield level and yield volatility for bonds with two-, five- and ten-year maturity (based on daily 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 1992:02–2010:12 there were 33 rate cuts, 31 hikes, and 4639 days on which the rate did not change. In panel *d* the dashed line represents conditional volatilities on the pre-crisis sample 1992:01–2007:12.

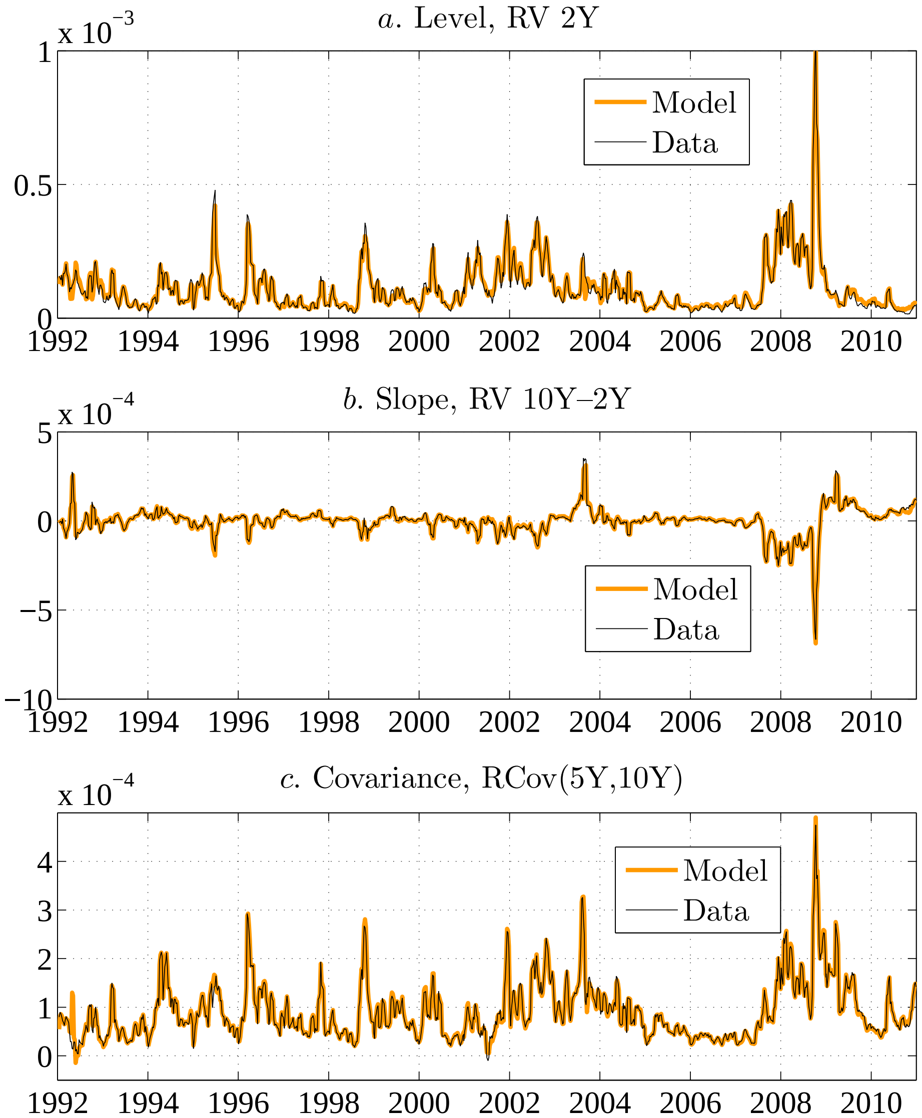


Figure 3.5: In-sample fit to the second moment dynamics of yields

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. The sample period is 1992:01 through 2010:12 with weekly sampling.

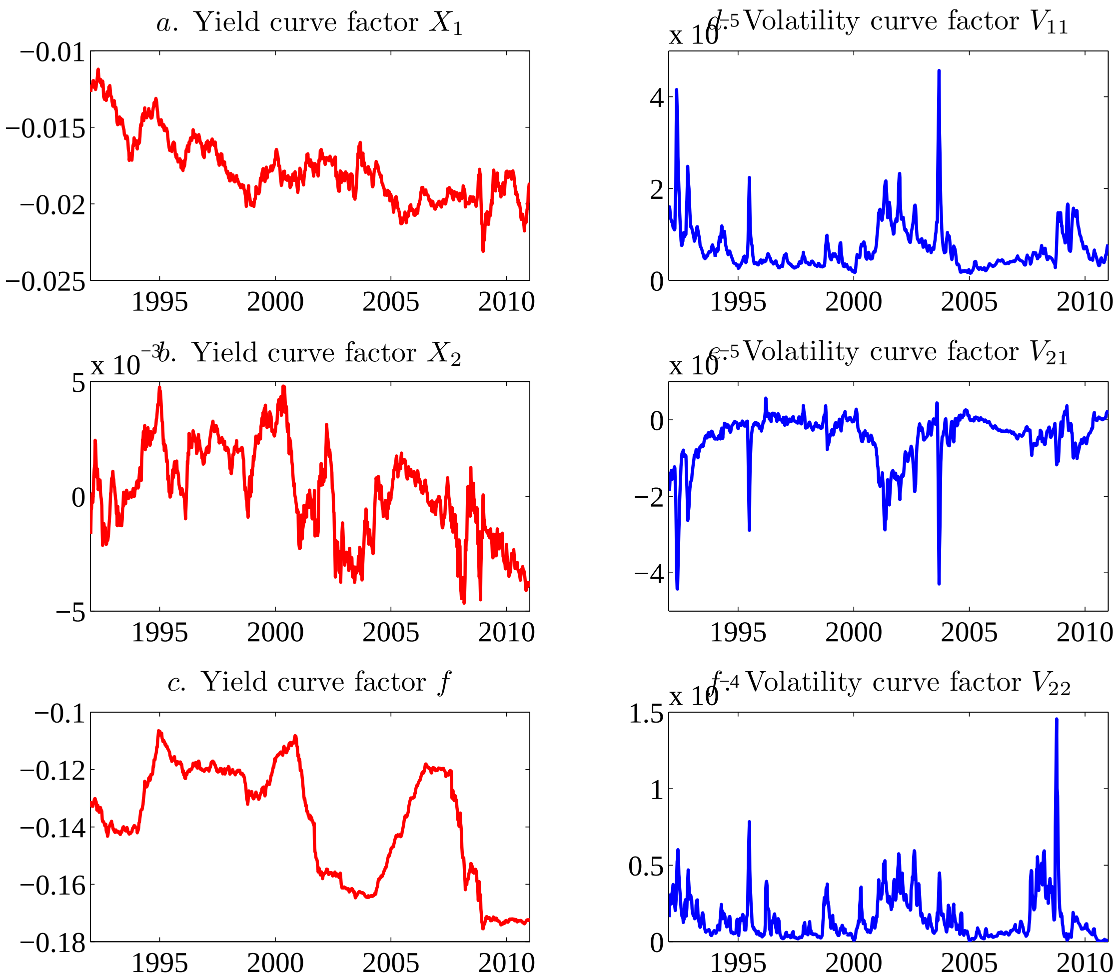


Figure 3.6: Factor dynamics filtered from the model

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. For presentation, the X states have been multiplied by -1. The sample period is 1992:01 through 2010:12 with weekly sampling.

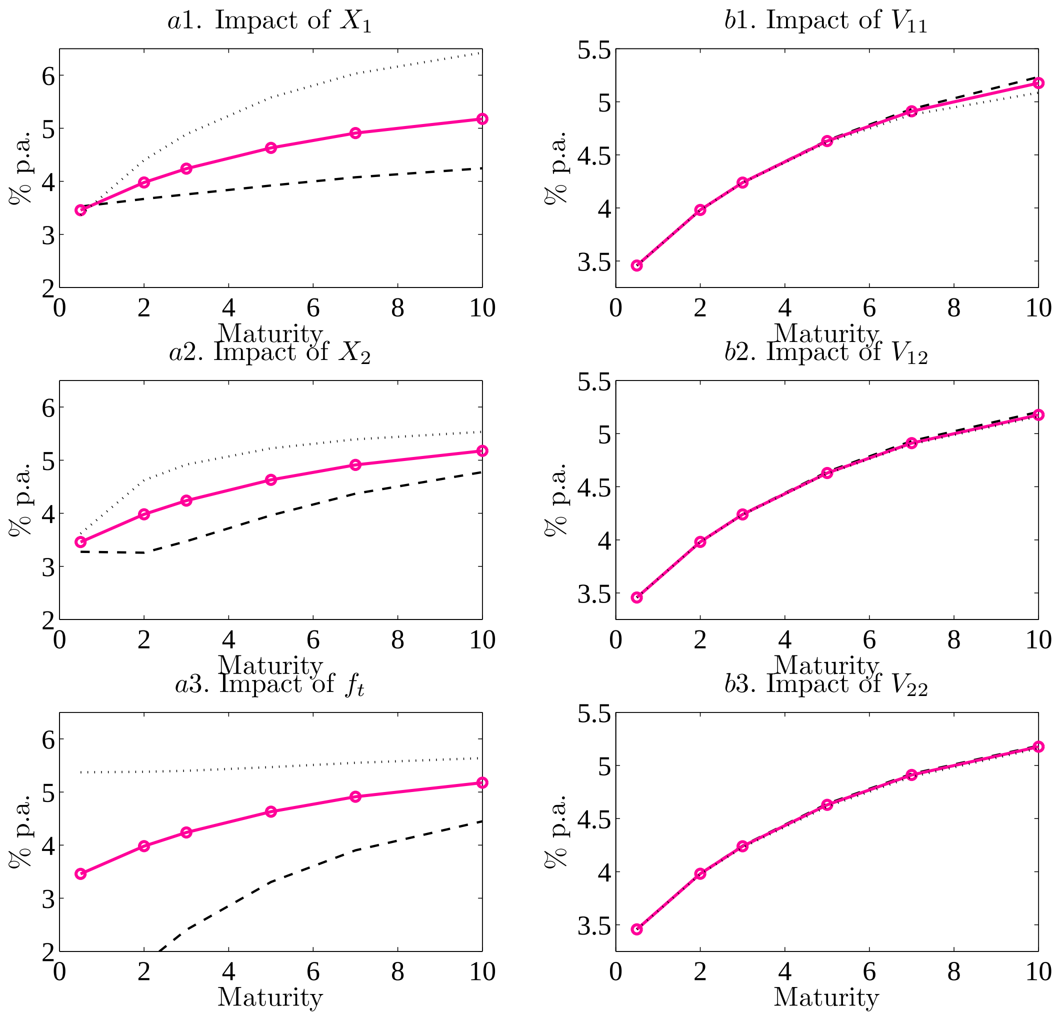


Figure 3.7: Factor shocks and yield curve responses

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. The sample period is 1992:01 through 2010:12 with weekly sampling.

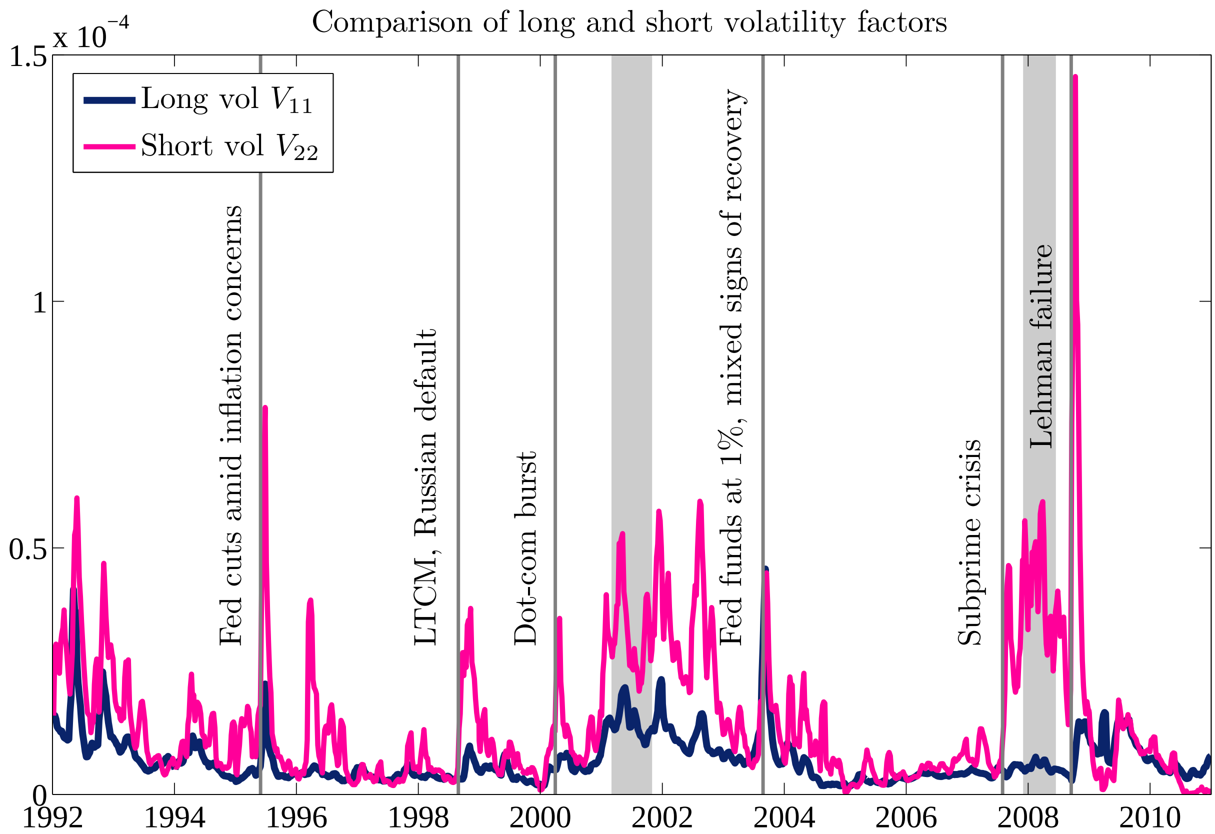


Figure 3.8: Short vs. long volatility factors and selected events

The figure plots the model-implied dynamics for volatility factors V_{11} and V_{22}. We discuss them in Section III.B. Shaded areas represent NBER-dated recessions. Vertical lines mark major economic or financial events that moved Treasury volatility. The sample period is 1992:01 through 2010:12 with weekly sampling.

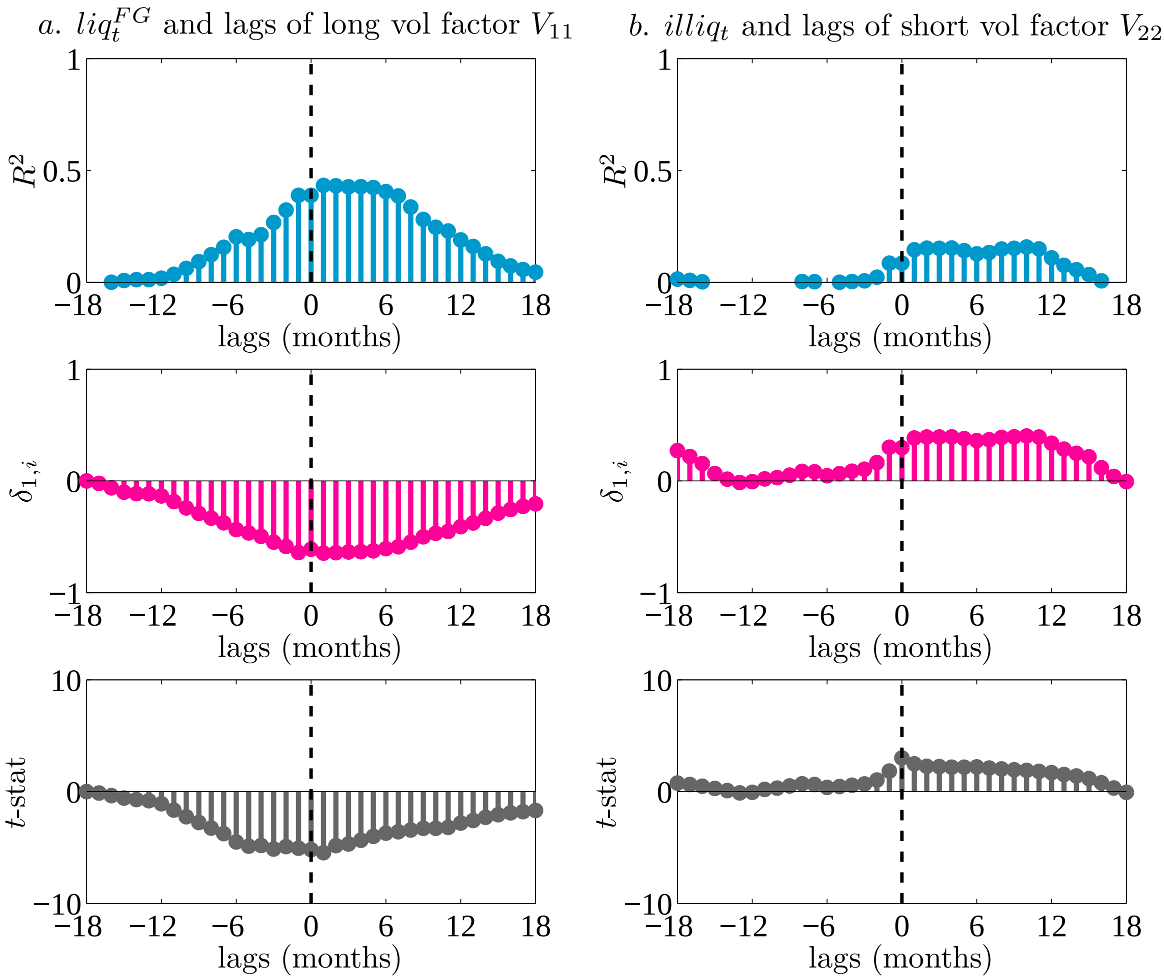


Figure 3.9: Liquidity and volatility factors, lead and lag relationship

The figure plots the lead-lag relationship between volatility and liquidity. Positive lags mean that volatility factors forecast liquidity. Panel a of the figure plots the R^2's, the regression coefficient $\delta_{i,1}$ and the t-statistic for the univariate regression of liquidity factor liq_t^{FG} on lagged long volatility state V_{11} using monthly data. Sample period is 1992:01 through 2008:12 and is determined by the availability of the Fontaine-Garcia liquidity factor. Panel b of the figure plots the R^2's, the regression coefficient $\delta_{i,1}$ and the t-statistic for the univariate regression of illiquidity factor iliq_t on lagged short volatility state V_{22} using monthly data. To obtain monthly data for the illiquidity factor, we take the average of daily data within the given month. Sample period is 1992:01 through 2009:12 and is determined by the availability of the illiquidity factor. In both panels, t-statistics are in parentheses and are computed using the Newey-West adjustment with 12 lags.

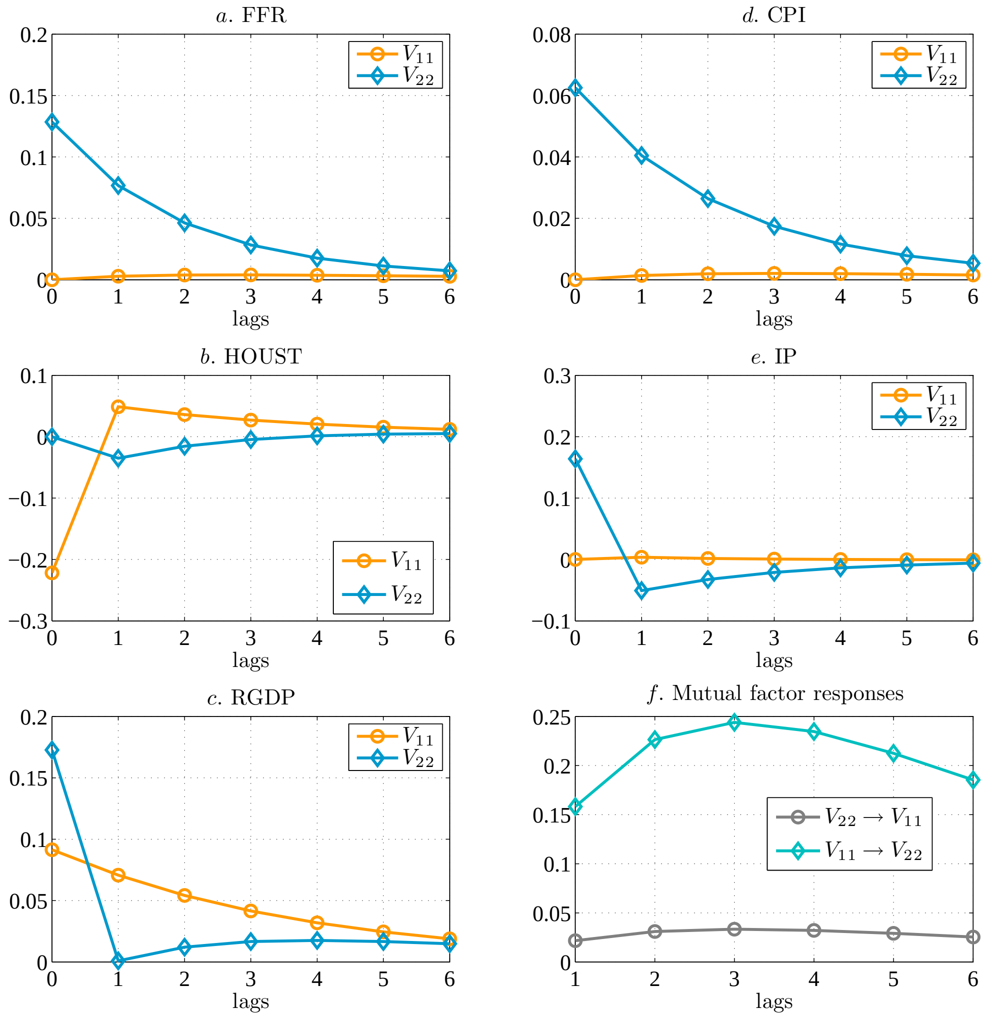


Figure 3.10: Impulse-response functions of volatility factors to macroeconomic uncertainty proxies

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 in the BlueChip survey. 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.

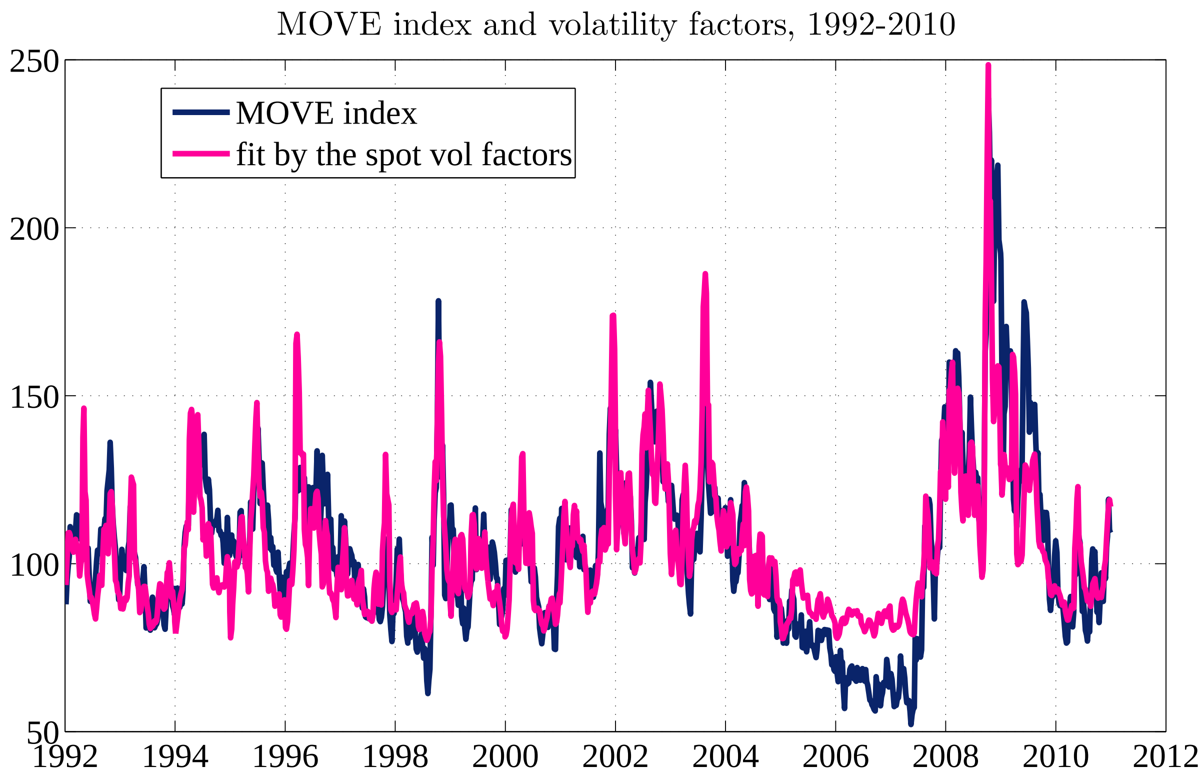


Figure 3.11: MOVE index and volatility factors

The figure plots the observed and fitted Merrill Lynch Option Volatility Estimate (MOVE) index. The fit is obtained from an OLS regression using three filtered volatility factors long, short volatility and the covolatility factor, V_{11}, V_{21}, V_{22}, respectively. Sample period is 1992:01 through 2010:12. The sampling frequency is weekly.

Table I: 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 1992:01–2010:12. JB denotes the Jarque-Bera normality test (critical value 5.975). 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	3.44	3.98	4.24	4.63	4.90	5.18
Skew	-0.42	-0.37	-0.36	-0.25	-0.09	0.13
Kurt	1.82	2.02	2.13	2.22	2.25	2.29
JB	87.66	63.38	53.24	35.88	25.06	23.99

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

	2Y	3Y	5Y	7Y	10Y
Mean	113.69	121.06	120.42	115.42	110.81
Stdev	43.77	45.11	38.88	34.43	32.47
Skew	1.38	1.19	0.88	0.75	1.01
Kurt	6.22	5.16	3.87	3.43	4.58
JB	754.27	432.95	160.54	102.62	274.56

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}	1.00										
y_t^{2Y}	0.97	1.00									
y_t^{3Y}	0.94	0.99	1.00								
y_t^{5Y}	0.87	0.96	0.98	1.00							
y_t^{7Y}	0.80	0.91	0.95	0.99	1.00						
y_t^{10Y}	0.72	0.85	0.90	0.96	0.99	1.00					
v_t^{2Y}	-0.16	-0.13	-0.11	-0.08	-0.05	*-0.02*	1.00				
v_t^{3Y}	-0.22	-0.17	-0.14	-0.09	*-0.05*	*0.00*	0.96	1.00			
v_t^{5Y}	-0.33	-0.29	-0.26	-0.22	-0.18	-0.14	0.89	0.90	1.00		
v_t^{7Y}	-0.40	-0.34	-0.31	-0.26	-0.22	-0.17	0.80	0.83	0.97	1.00	
v_t^{10Y}	-0.37	-0.31	-0.27	-0.22	-0.17	-0.12	0.72	0.76	0.89	0.93	1.00

Table II: Parameter estimates

This table reports parameter estimates for two models: "G_3SV_0" is the Gaussian three-factor model with essentially affine market prices of risk; "G_3SV_3" denotes a joint model of yield and volatility curve with three conditionally Gaussian factors and three volatility factors. "G_3SV_3 1992-2007" indicates the parameter estimates obtained on the sample 1992:01–2007:12. The Gaussian model follows Duffee (2002) using the normalization of Dai and Singleton (2000), which 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 divided by the number of observations T. For comparison, in models with stochastic volatility, we split the log-likelihood values into the yield and the volatility components: Loglik y_t^τ/T and Loglik v_t^τ/T. The last column reports log-likelihood values obtained from parameter estimates on the 1992:01–2007:12 sample and applied to the 1992:01–2010:12 sample.

	G_3SV_0	G_3SV_3	G_3SV_3 1992-2007
$\mathcal{K}_{11}$	-0.141	-0.020	-0.033
	(0.091)	(0.021)	(0.008)
$\mathcal{K}_{21}$	-1.735	–	–
	(0.067)	–	–
$\mathcal{K}_{22}$	-0.558	-1.104	-1.947
	(0.087)	(0.143)	(0.309)
$\mathcal{K}_{31}$	-1.919	-3.073	-2.768
	(0.074)	(0.191)	(0.175)
$\mathcal{K}_{32}$	-0.889	-7.117	-7.668
	(0.087)	(0.437)	(0.479)
$\mathcal{K}_{33}$	-0.001	-0.460	-0.498
	(0.002)	(0.010)	(0.008)
σ_f	–	0.006	0.006
	–	(7.62e-5)	(7.05e-5)
k	–	2.000	2.000
	–	(0.239)	(1.067)
Q_{11}	–	0.001	0.001
	–	(9.63e-5)	(1.14e-4)
Q_{22}	–	0.003	0.003
	–	(2.59e-4)	(2.63e-4)
M_{11}	–	-0.042	-0.113
	–	(0.035)	(0.038)
M_{21}	–	-0.205	-0.185
	–	(0.179)	(0.177)
M_{22}	–	-0.599	-0.685
	–	(0.091)	(0.127)
γ_0	1.000	0.153	0.159
	(1.550)	(0.089)	(0.026)
γ_{11}	0.001	–	–
	(2.63e-4)	–	–
γ_{12}	0.004	–	–
	(1.60e-4)	–	–
γ_{13}	0.005	–	–
	(1.01e-4)	–	–
λ_{11}^{1X}	0.460	0.004	-0.018
	(0.068)	(0.021)	(0.020)
λ_{22}^{1X}	0.222	-0.054	-0.703
	(0.066)	(0.144)	(0.310)
σ_{MY}	1.58e-7	1.37e-7	7.84e-8
	(2.16e-9)	(1.75e-9)	(1.40e-9)
Loglik y_t^τ/T	40.632	40.744	40.191
Loglik v_t^τ/T	–	31.063	30.989

Table III: Model fit

This table gives the in-sample fit for three model specifications presented in Table II. The performance of our model (G_3SV_3) in matching yields is compared with the Gaussian setting (G_3SV_0). The column "G_3SV_3 1992-2007" reports the model fit for the period 1992–2010 using parameter estimated on the 1992:01–2007:12 sample. 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	G_3SV_3	G_3SV_3 1992-2007
Panel a. Term structure of yields			
$a1$. RMSE in bps			
6 month	2.09	2.02	1.63
2 year	3.70	3.04	3.28
3 year	2.59	2.63	2.80
5 year	3.89	3.50	3.50
7 year	2.56	2.27	2.13
10 year	3.65	3.34	3.46
$a2$. Explained variation (%)			
6 months	99.99	99.99	99.99
2 year	99.97	99.97	99.98
3 year	99.99	99.98	99.98
5 year	99.94	99.95	99.95
7 year	99.97	99.97	99.98
10 year	99.93	99.93	99.93
Panel b. Interest rate volatilities			
$b1$. RMSE in bps			
v_t^{2Y}	–	5.72	6.10
$v_t^{10Y} - v_t^{2Y}$	–	5.22	6.09
$v_t^{5Y,10Y}$	–	4.06	4.41
$b2$. Explained variation (%)			
v_t^{2Y}	–	98.25	97.99
$v_t^{10Y} - v_t^{2Y}$	–	99.15	98.66
$v_t^{5Y,10Y}$	–	98.33	98.00

Table IV: 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. $f, X_1, X_2, V_{11}, V_{12}, V_{22}$. The sample covers the 1992:01–2010:12 period, with weekly sampling frequency. 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.	3.69	2.37	0.56	1.71	1.45	3.25	2.80	0.49
X_1	1.07	-1.93	0.94	-1.11	1.48	0.16	-0.89	-2.91
X_2	-0.34	1.03	0.34	1.81	-0.00	0.54	-2.15	0.75
f	4.25	1.26	1.22	0.97	4.07	3.03	1.66	-2.91
V_{11}	8.10	3.93	13.50	16.68	28.86	16.79	13.00	30.10
V_{21}	34.84	7.79	21.62	31.51	39.83	36.60	30.18	89.30
V_{22}	109.50	54.06	29.97	31.84	25.80	66.90	31.48	69.57
$\bar{R}^2$	0.99	0.92	0.91	0.93	0.95	0.97	0.95	0.98

Table V: Interest rate volatility and liquidity

Panel a reports the regression results of noise illiquidity measure of Hu, Pan, and Wang (2013) on volatility factors V_{11}, V_{21}, V_{22}. The weekly data for illiquidity measure are obtained by taking the average of daily data within the week. Panel b reports the regression results of weekly settlement failures in Treasury bonds on volatility factors. All variables are standardized. t-statistics are in parentheses and are computed using the Newey-West adjustment with 12 lags.

Panel a. Noise illiquidity measure
$\text{illiq}_t = a_0 + a_1 V_{11t} + a_2 V_{21t} + a_3 V_{22t} + \varepsilon_t$

V_{11}	0.21	0.10	0.66
	(2.72)	(0.90)	(2.09)
V_{21}	–	0.24	0.30
	(–)	(1.74)	(2.97)
V_{22}	–	–	0.71
	(–)	(–)	(2.24)
$\bar{R}^2$	0.05	0.09	0.25

Panel b. Failure to deliver
$\text{fails}_t = a_0 + a_1 V_{11t} + a_2 V_{21t} + a_3 V_{22t} + \varepsilon_t$

V_{11}	0.12	-0.13	0.35
	(1.27)	(-0.65)	(1.33)
V_{21}	–	0.49	0.55
	(–)	(1.83)	(2.22)
V_{22}	–	–	0.60
	(–)	(–)	(2.53)
$\bar{R}^2$	0.01	0.19	0.31

Table VI: 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 1992:01–2010:12. This amounts to 228 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)	0.124	–	–	–	-0.402	–
	(0.054)	–	–	–	(0.068)	–
σ(RGDP)	–	–	–	–	–	–
	–	–	–	–	–	–
E(UNEMPL)	0.156	–	–	–	–	–
	(0.051)	–	–	–	–	–
σ(UNEMPL)	–	–	–	0.251	-0.387	–
	–	–	–	(0.095)	(0.094)	–
E(HOUST)	–	–	–	–	–	–
	–	–	–	–	–	–
σ(HOUST)	–	-0.190	–	-0.181	–	–
	–	(0.068)	–	(0.088)	–	–
E(FFR)	–	0.765	0.950	-0.318	–	–
	–	(0.088)	(0.018)	(0.121)	–	–
σ(FFR)	–	-0.141	–	–	-0.261	0.354
	–	(0.064)	–	–	(0.081)	(0.081)
E(IP)	–	–	0.044	–	–	–
	–	–	(0.020)	–	–	–
σ(IP)	–	0.253	–	–	–	0.207
	–	(0.068)	–	–	–	(0.100)
E(CPI)	0.888	–	0.074	–	–	–
	(0.055)	–	(0.017)	–	–	–
σ(CPI)	–	–	–	–	0.194	0.263
	–	–	–	–	(0.089)	(0.158)
$\bar{R}^2$	0.73	0.62	0.98	0.16	0.26	0.20

VIII. Appendix

VIII.A. Data description

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

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

Table VII: Predictability of bond excess returns and volatility

The table reports the results obtained from the predictive regression given in equation (3.23). Panel a reports the restricted regression with only $\widehat{cf}_t$ as predictor. Panel b reports the regression result when two volatility factors are used as regressors, finally panel c uses $\widehat{cf}_t$, V_{11} and V_{22} as return predictors. ΔR^2 V_{11} denotes the increase in R^2's obtained by including V_{11} next to $\widehat{cf}_t$ in the regression, ΔR^2 V_{22} denotes the marginal increase in R^2's by including V_{22} next to $\widehat{cf}_t$. The data are sampled monthly and the sample period is 1992:01–2010:12. In parentheses, we report t-statistics obtained by a Newey-West adjustment with 12 lags.

Panel a. $rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 \widehat{cf}_t + \varepsilon_{t+1}^{(n)}$

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
$\widehat{cf}$	0.43	0.49	0.58	0.65	0.69
	(3.17)	(3.59)	(4.71)	(5.76)	(6.36)
$\bar{R}^2$	0.18	0.23	0.34	0.43	0.47

Panel b. $rx_{t+1}^{(n)} = \gamma_0 + \gamma_2 V_{11t} + \gamma_3 V_{22t} + \varepsilon_{t+1}^{(n)}$

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
V_{11}	0.11	0.13	0.15	0.16	0.11
	(0.90)	(1.03)	(1.32)	(1.35)	(0.92)
V_{22}	0.28	0.25	0.22	0.16	0.18
	(2.78)	(2.35)	(2.10)	(1.60)	(1.79)
$\bar{R}^2$	0.13	0.11	0.11	0.08	0.06

Panel c. $rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 \widehat{cf}_t + \gamma_2 V_{11t} + \gamma_3 V_{22t} + \varepsilon_{t+1}^{(n)}$

	$rx^{(2)}$	$rx^{(3)}$	$rx^{(5)}$	$rx^{(7)}$	$rx^{(10)}$
$\widehat{cf}$	0.39	0.46	0.56	0.64	0.69
	(3.13)	(3.59)	(4.84)	(5.96)	(6.78)
V_{11}	0.00	-0.00	-0.00	-0.02	-0.09
	(0.02)	(-0.02)	(-0.02)	(-0.21)	(-0.78)
V_{22}	0.31	0.27	0.25	0.20	0.22
	(2.84)	(2.37)	(2.12)	(1.70)	(1.86)
$\bar{R}^2$	0.27	0.30	0.40	0.46	0.50
$\Delta\bar{R}^2$ V_{11}	0.03	0.02	0.02	0.01	0.00
$\Delta\bar{R}^2$ V_{22}	0.06	0.05	0.04	0.02	0.03

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. 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 (2010) and delete days with no trading for more than three hours.

We choose the ten-minute 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 Figure 3.12.

Table VIII: Option implied versus spot yield volatility

The table reports the regression results of MOVE index on volatility factors V_{11}, V_{21} and V_{22} using weekly data. Weekly MOVE index is obtained by taking the average of daily data within given week. All variables are standardized. T-statistics are in parentheses and are computed using the Newey-West adjustment with 12 lags.

$$MOVE_t = a_0 + a_1 V_{11t} + a_2 V_{21t} + a_3 V_{22t} + a_4 VIX_t + \varepsilon_t$$

V_{11}	0.34	0.07	0.92	0.79
	(3.70)	(0.76)	(6.01)	(6.82)
V_{21}	–	0.52	0.62	0.52
	(–)	(5.99)	(16.11)	(10.22)
V_{22}	–	–	1.08	0.91
	(–)	(–)	(7.92)	(9.39)
VIX	–	–	–	0.23
	(–)	(–)	(–)	(3.77)
$\bar{R}^2$	0.11	0.32	0.68	0.72

The liquidity in the secondary bond market is concentrated in two-, three-, five- and ten-year securities (see also Fleming and Mizrach, 2009, Table 1). We assume that the dynamics of this most liquid segment spans the information content of the whole curve. Since any method for bootstrapping the zero curve 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.

GovPX and BrokerTec

Table IX reports some basic statistics on the Treasury bond transaction data in the period 1992:01 through 2010:12. For the GovPX period (1992:01–2000:12), we report the average number of quotes per trading day and for the BrokerTec period (2001:01–2010:12) we report the average number of transactions per trading day. The number of Treasury bonds and bills totals to 1148 in our sample period. These were transacted or quoted more than 49.3 million times in the on-the-run secondary market.

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 X 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

Table IX: 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	1414
3Y	1385	1017
5Y	3128	2801
7Y	637	1500
10Y	2649	2659
30Y	793	1114

three, five and ten years. However, the magnitude of all autocorrelations is very small, which makes them economically insignificant.

Table X: 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 (Figure 3.12). 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.

Realized covariance matrix estimation

This section discusses the robustness and efficiency of the realized second moment estimator proposed in the paper that are critical for our results. In the paper, we use the outer product estimator given by:

$$RCov(t, t+1; N) = \sum_{i=1,\ldots,N} \left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}}\right)\left(y_{t+\frac{i}{N}} - y_{t+\frac{i-1}{N}}\right)'. \tag{3.25}$$

In case of asynchronous trading, the realized covariance matrix estimator defined in Eq. (3.25) 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

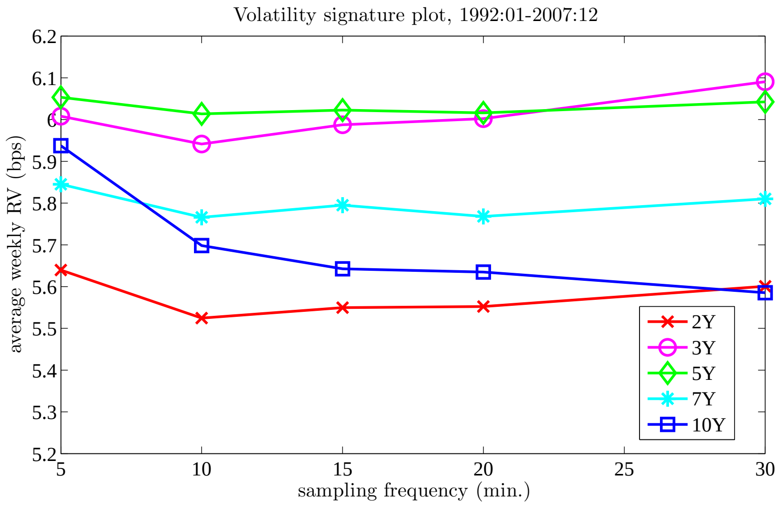


Figure 3.12: Volatility signature plot

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.

depends on the difference in liquidity of the assets considered.[25] Hayashi and Yoshida (2005, HY) propose a covariance estimator which corrects for the bias in (3.25). 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:

$$RCov_{i,j}^{HY}(t, t+h; N_i, N_j) = \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \left(y_{t_i + \frac{kh}{N_i}} - y_{t_i + \frac{(k-1)h}{N_i}} \right) \left(y_{t_j + \frac{lh}{N_j}} - y_{t_j + \frac{(l-1)h}{N_j}} \right) \mathbb{I}\left(\tau_i \cap \tau_j \neq \emptyset\right) \quad (3.26)$$

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 (3.25), 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 (3.25). 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

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

Table XI: 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

liquidity, which is well proxied by the average number of quotes/trades per day reported in Table IX of Appendix VIII.A.

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 very long end of the curve. In the period 2001:01–2010:12, 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 Gürkaynak, Sack, and Wright (2006) (GSW). Our daily yields are almost perfectly correlated with the CMTs as well as with the GSW yields. Table XI summarizes the results.

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.

VIII.B. Model solution

We provide solutions 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.

Dependence between X and V factors

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

$$dZ_X = dW\rho + \sqrt{1-\rho'\rho}\; dB \tag{3.27}$$

$$= dW\rho + \widetilde{\rho} dB, \tag{3.28}$$

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, 2007; Buraschi, Porchia, and Trojani, 2010). We use short notation $\widetilde{\rho} := \sqrt{1-\rho'\rho}$.

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 $\mathbb{P}$):

$$dZ = \begin{pmatrix} dZ_X \\ dZ_f \end{pmatrix} = \begin{pmatrix} dW\rho + \widetilde{\rho}dB \\ dZ_f \end{pmatrix} = \begin{pmatrix} dW\rho \\ 0_{1\times 2} \end{pmatrix} + \underbrace{\begin{pmatrix} \widetilde{\rho}I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & 1_{1\times 1} \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} dB \\ dZ_f \end{pmatrix}}_{d\widetilde{Z}} = \begin{pmatrix} dW\rho \\ 0_{1\times 2} \end{pmatrix} + Rd\widetilde{Z}, \tag{3.29}$$

where

$$R = \begin{pmatrix} \widetilde{\rho}I_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 2} & 1_{1\times 1} \end{pmatrix}, \qquad d\widetilde{Z} = \begin{pmatrix} dB \\ dZ_f \end{pmatrix}. \tag{3.30}$$

The change of drift is specified as:

$$d\widetilde{Z} = d\widetilde{Z}^{\mathbb{Q}} - \Lambda_{Y,t}dt \tag{3.31}$$

$$dW = dW^{\mathbb{Q}} - \Lambda_{V,t}dt \tag{3.32}$$

$$\Lambda_{Y,t} = \Sigma_Y^{-1}(V_t)\left(\lambda_Y^0 + \lambda_Y^1 Y_t\right) \tag{3.33}$$

$$\Lambda_{V,t} = \left(\sqrt{V_t}\right)^{-1}\Lambda_V^0 + \sqrt{V_t}\Lambda_V^1, \tag{3.34}$$

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:

$$dY_t = \left[\left(\mu_Y - \begin{pmatrix} \Lambda_V^0\rho \\ 0 \end{pmatrix} - R\lambda_Y^0\right) + \left(\mathcal{K}_Y - R\lambda_Y^1\right)Y_t - \begin{pmatrix} V_t\Lambda_V^1\rho \\ 0 \end{pmatrix}\right] dt + \Sigma_Y\left(V_t\right) dZ_t^{\mathbb{Q}}. \tag{3.35}$$

Let:

$$\mu_Y^Q = \mu_Y - \left(\begin{array}{c} \Lambda_V^0 \rho \\ 0 \end{array} \right) - R\lambda_Y^0 \tag{3.36}$$

$$\mathcal{K}_Y^Q = \mathcal{K}_Y - R\lambda_Y^1. \tag{3.37}$$

The dynamics of V_t is given as:

$$dV_t = \left[\left(\Omega\Omega' - \Lambda_V^0 Q - Q'\Lambda_V^{0\prime} \right) + \left(M - Q'\Lambda_V^{1\prime} \right) V_t + V_t \left(M' - \Lambda_V^1 Q \right) \right] dt + \sqrt{V_t} dW_t^{\mathbb{Q}} Q + Q' dW_t^{\mathbb{Q}\prime} \sqrt{V_t}. \tag{3.38}$$

Let

$$\Omega^{\mathbb{Q}}\Omega^{\mathbb{Q}\prime} = \Omega\Omega' - \Lambda_V^0 Q - Q'\Lambda_V^{0\prime} = (k - 2v) Q'Q \tag{3.39}$$

$$M^{\mathbb{Q}} = M - Q'\Lambda_V^{1\prime}, \tag{3.40}$$

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

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:

$$F(Y_t, V_t; t, \tau) = \exp\left\{ A(\tau) + B(\tau)' Y_t + Tr\left[C(\tau) V_t \right] \right\}. \tag{3.41}$$

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

$$\mathcal{L}_{\{Y,V\}} F + \frac{\partial F}{\partial t} = rF, \tag{3.42}$$

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

$$\mathcal{L}_{\{Y,V\}} F = \left(\mathcal{L}_Y + \mathcal{L}_V + \mathcal{L}_{Y,V} \right) F \tag{3.43}$$

$$\mathcal{L}_Y F = \frac{\partial F}{\partial Y'} \left[\mu_Y^{\mathbb{Q}} + \mathcal{K}_Y^{\mathbb{Q}} Y - \left(\begin{array}{c} V_t \Lambda_V^1 \rho \\ 0 \end{array} \right) \right] + \frac{1}{2} Tr \left[\frac{\partial F}{\partial Y \partial Y'} \Sigma_Y(V) \Sigma_Y'(V) \right] \tag{3.44}$$

$$\mathcal{L}_V F = Tr \left[\left(\Omega^{\mathbb{Q}}\Omega^{\mathbb{Q}\prime} + M^{\mathbb{Q}} V + V M^{\mathbb{Q}\prime} \right) \mathcal{R} F + 2V\mathcal{R}Q'Q\mathcal{R}F \right] \tag{3.45}$$

$$\mathcal{L}_{Y,V} F = 2Tr \left[\left(\mathcal{R} Q' \rho \frac{\partial}{\partial X'} \right) FV \right]. \tag{3.46}$$

$\mathcal{R}$ is a matrix differential operator: $\mathcal{R}_{ij} := \left(\frac{\partial}{\partial V_{ij}}\right)$. Substituting derivatives of (3.41) into (3.42) gives:

$$B_\tau' \left(\mu_Y^{\mathbb{Q}} + \mathcal{K}_Y^{\mathbb{Q}} Y\right) - Tr\left(\Lambda_V^1 \rho B_{X,\tau}' V\right) + \frac{1}{2} Tr\left(B_{X,\tau} B_{X,\tau}' V\right) + \frac{1}{2} B_{f,\tau}^2 \sigma_f^2 \tag{3.47}$$

$$+ Tr\left(\Omega^{\mathbb{Q}} \Omega^{\mathbb{Q}\prime} C_\tau\right) + Tr\left[\left(C_\tau M^{\mathbb{Q}} + M^{\mathbb{Q}\prime} C_\tau + 2C_\tau Q' Q C_\tau\right) V\right] \tag{3.48}$$

$$+ Tr\left[\left(C_\tau Q' \rho B_{X,\tau}' + B_{X,\tau} \rho' Q C_\tau\right) V\right] \tag{3.49}$$

$$= \frac{\partial A_\tau}{\partial \tau} + \frac{\partial B_\tau}{\partial \tau}' Y + Tr\left(\frac{\partial C_\tau}{\partial \tau} V\right) + \gamma_0 + \gamma_Y' Y \tag{3.50}$$

By matching coefficients, we obtain the system of equations:

$$\frac{\partial A}{\partial \tau} = B_\tau' \mu_Y^{\mathbb{Q}} + \frac{1}{2} B_{f,\tau}^2 \sigma_f^2 + Tr\left(\Omega^Q \Omega^{Q\prime} C_\tau\right) - \gamma_0 \tag{3.51}$$

$$\frac{\partial B}{\partial \tau} = \mathcal{K}^{\mathbb{Q}\prime} B_\tau - \gamma_Y \tag{3.52}$$

$$\frac{\partial C}{\partial \tau} = \frac{1}{2} B_{X,\tau} B_{X,\tau}' + C_\tau \left(M^{\mathbb{Q}} + Q' \rho B_{X,\tau}'\right) + \left(M^{\mathbb{Q}\prime} + B_{X,\tau} \rho' Q\right) C_\tau + 2C_\tau Q' Q C_\tau - \Lambda_V^1 \rho B_{X,\tau}' \tag{3.53}$$

To obtain the solution provided in the text, set $\rho = 0_{2\times1}$, $\Lambda_V^0 = 0_{2\times2}$ and $\Lambda_V^1 = 0_{2\times2}$.

Instantaneous volatility of yields

The instantaneous volatility of yields is given as:

$$\frac{1}{dt} \langle dy_t^{\tau_1}, dy_t^{\tau_2} \rangle = \frac{1}{\tau_1 \tau_2} Tr\left[B_{f,\tau_2} B_{f,\tau_1} \sigma_f^2 + \left(B_{X,\tau_1} B_{X,\tau_2}' + 2C_{\tau_2} Q' \rho B_{X,\tau_1}' + 2C_{\tau_1} Q' \rho B_{X,\tau_2}' + 4C_{\tau_1} Q' Q C_{\tau_2}\right) V_t\right]. \tag{3.54}$$

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

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

Where we use the following fact:

$$Tr\left[C\left(\sqrt{V} dW Q + Q' dW' \sqrt{V}\right)\right] = 2Tr\left(Q C \sqrt{V} dW\right). \tag{3.55}$$

$\square$

Conditional covariance of X and V

We consider the conditional covariance matrix of X and V

$$d\left\langle \begin{pmatrix} X_{t,1} \\ X_{t,2} \\ f_t \end{pmatrix}, \begin{pmatrix} V_{t,11} \\ V_{t,12} \\ V_{t,22} \end{pmatrix} \right\rangle = \begin{pmatrix} d\left(X_1, V_{11}\right) & d\left(X_1, V_{12}\right) & d\left(X_1, V_{22}\right) \\ d\left(X_2, V_{11}\right) & d\left(X_2, V_{12}\right) & d\left(X_2, V_{22}\right) \\ d\left(f, V_{11}\right) & d\left(f, V_{12}\right) & d\left(f, V_{22}\right) \end{pmatrix} \tag{3.56}$$

The elements of the covariance matrix are given by:

$$d\left\langle X_k, V_{ij}\right\rangle = \rho'\left(Q_{:,j}V_{ik} + Q_{:,i}V_{jk}\right), \tag{3.57}$$

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

Proof. The expression follows by simple algebra:

$$\begin{aligned} \frac{1}{dt} d\left\langle V_{ij}, X_k\right\rangle &= \left[e_i'\left(\sqrt{V}dWQ\right)e_j + e_i'\left(Q'dW'\sqrt{V}\right)e_j\right]\left(e_k'\sqrt{V}dW\rho\right) \\ &= Tr\left(e_je_i'\sqrt{V}dWQ\right) \times Tr\left(\rho e_k'\sqrt{V}dW\right) + Tr\left(e_je_i'Q'dW'\sqrt{V}\right) \times Tr\left(\rho e_k'\sqrt{V}dW\right) \\ &= \text{vec}\left(\sqrt{V}e_ie_j'Q'\right)'\text{vec}\left(\sqrt{V}e_k\rho'\right) + \text{vec}\left(\sqrt{V}e_je_i'Q'\right)'\text{vec}\left(\sqrt{V}e_k\rho'\right) \\ &= Tr\left(Qe_je_i'Ve_k\rho'\right) + Tr\left(Qe_ie_j'Ve_k\rho'\right) \\ &= \rho'\left(Q_{:,j}V_{ik} + Q_{:,i}V_{jk}\right), \end{aligned} \tag{3.58}$$

where e_i is the i-th column of the identity matrix. $\square$

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:

$$X_{t+\Delta t} = \bar{\mu}_{X,\Delta t} + \Phi_{X,\Delta t}X_t + \sqrt{V_t\Delta t}\left(U_{t+\Delta t}\rho + \widetilde{\rho}b_{t+\Delta t}\right) \tag{3.59}$$

$$V_{t+\Delta t} = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}' + \sqrt{V_t\Delta t}U_{t+\Delta t}Q + Q'U_{t+\Delta t}'\sqrt{V_t\Delta t}, \tag{3.60}$$

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 :

$$Cov\left[X, \text{vec}(V)\right] = E\left[X\text{vec}\left(V\right)'\right] - E\left(X\right)E\left[\text{vec}\left(V\right)'\right]. \tag{3.61}$$

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

$$\mathrm{vec}E\left[X\left(\mathrm{vec}V\right)'\right]=\left[I_{n^3}-\left(\Phi_V\otimes\Phi_V\right)\otimes\Phi_X\right]^{-1}\left(\mathrm{vec}A+\mathrm{vec}B\right), \tag{3.62}$$

where A is given as:

$$A=\bar{\mu}_X\mathrm{vec}\left(k\bar{\mu}_V\right)'+\bar{\mu}_X\mathrm{vec}\left[\Phi_V E\left(V_t\right)\Phi_V'\right]+\Phi_X E(X_t)\mathrm{vec}(k\bar{\mu}_V)', \tag{3.63}$$

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

$$B_{k,ij}=\rho'\left(Q_{:,j}V_{ik}+Q_{:,i}V_{jk}\right)\Delta t, \tag{3.64}$$

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.

VIII.C. Moments of the state variables

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

Moments of the V_t process

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

$$E_t\left(V_{t+\Delta t}\right)=k\bar{\mu}_{V,\Delta t}+\Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}', \tag{3.65}$$

where

$$\Phi_{V,\Delta t}=e^{M\Delta t} \tag{3.66}$$

$$\bar{\mu}_{V,\Delta t}=\int_0^{\Delta t}e^{M\Delta t}Q'Qe^{M'\Delta t}ds=-\frac{1}{2}\hat{C}_{12}(\Delta t)\hat{C}_{11}'\left(\Delta t\right), \tag{3.67}$$

with

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

See Van Loan (1978) for the derivation of the above expression. Assuming stationarity (i.e. negative eigenvalues of M), the unconditional first moment of V_t follows as:

$$\lim_{\Delta t \to \infty} \text{vec} E_t \left(V_{t+\Delta t} \right) = k \text{vec} \left(\bar{\mu}_{V,\infty} \right) = -k \left[(I \otimes M) + (M \otimes I) \right]^{-1} \text{vec}(Q'Q). \tag{3.68}$$

The conditional and unconditional covariance matrix of V_t reads:

$$Cov_t \left[\text{vec} \left(V_{t+\Delta t} \right) \right] = \left(I_{n^2} + K_{n,n} \right) \left[\Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} \otimes \bar{\mu}_{V,\Delta t} + k \left(\bar{\mu}_{V,\Delta t} \otimes \bar{\mu}_{V,\Delta t} \right) + \bar{\mu}_{V,\Delta t} \otimes \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} \right]. \tag{3.69}$$

$$\lim_{\Delta t \to \infty} Cov_t \left[\text{vec} \left(V_{t+\Delta t} \right) \right] = \left(I_{n^2} + K_{n,n} \right) k \left(\bar{\mu}_{V,\infty} \otimes \bar{\mu}_{V,\infty} \right). \tag{3.70}$$

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

Gourieroux, Jasiak, and Sufana (2009b) 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:

$$V_t = \sum_{i=1}^{k} v_t^i v_t^{i\prime} \tag{3.71}$$

$$v_{t+\Delta t}^i = \Phi_{V,\Delta t} v_t^i + \epsilon_{t+\Delta t}^i, \quad \epsilon_t^i \sim N(0, \bar{\mu}_{V,\Delta t}). \tag{3.72}$$

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

$$V_{t+\Delta t} = k\bar{\mu}_{V,\Delta t} + \Phi_{V,\Delta t} V_t \Phi'_{V,\Delta t} + u_{t+\Delta t}^V, \tag{3.73}$$

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

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 $Y_t = (X_t', f_t)'$:

$$dY_t = \left(\mu_Y + \mathcal{K}_Y Y_t \right) dt + \Sigma \left(V_t \right) dZ_t. \tag{3.74}$$

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

$$E_t \left(Y_{t+\Delta t} \right) = \left(e^{\mathcal{K}_Y \Delta t} - I \right) \mathcal{K}_Y^{-1} \mu_Y + e^{\mathcal{K}_Y \Delta t} Y_t \tag{3.75}$$

$$\lim_{\Delta t \to \infty} E_t \left(Y_{t+\Delta t} \right) = -\mathcal{K}_Y^{-1} \mu_Y, \tag{3.76}$$

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

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

$$d\hat{Y}(t,T) = \hat{\sigma}_Y(t,T)\, dZ_t, \tag{3.77}$$

where $\hat{\sigma}_Y(t,T) := \Phi_Y(t,T)\,\Sigma(V_t)$, with

$$\Phi_Y(t,T) = e^{\mathcal{K}_Y(T-t)} \tag{3.78}$$

and

$$\Sigma_Y(V_t) = \begin{pmatrix} \sqrt{V_t} & 0 \\ 0 & \sigma_f^2 \end{pmatrix}. \tag{3.79}$$

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

$$Y_T = \hat{Y}_{T,T} = \hat{Y}_{t,T} + \int_{s=t}^{T} \hat{\sigma}_Y(s,T)\, dZ_s. \tag{3.80}$$

Therefore, we have:

$$V_Y(t,T) = Cov_t\left[\int_{s=t}^{T} \hat{\sigma}_Y(s,T) dZ_s^Y\right] = E_t\left[\int_{s=t}^{T} \hat{\sigma}_Y(s,T)\hat{\sigma}_Y(s,T)' ds\right] \tag{3.81}$$

$$= \int_{s=t}^{T} \Phi_Y(s,T) E_t \begin{pmatrix} V_s & 0 \\ 0 & \sigma_f^2 \end{pmatrix} \Phi_Y'(s,T) ds. \tag{3.82}$$

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:

$$\Phi_Y(t,T) = \begin{pmatrix} \Phi_X(t,T) & 0 \\ \Phi_{Xf}(t,T) & \Phi_f(t,T) \end{pmatrix}. \tag{3.83}$$

Let us for convenience define two matrices:

$$\mathcal{M}_{1Y}(t,T) = \begin{pmatrix} \Phi_X(t,T) \otimes \Phi_X(t,T) \\ \Phi_X(t,T) \otimes \Phi_{fX}(t,T) \\ \Phi_{fX}(t,T) \otimes \Phi_X(t,T) \\ \Phi_{fX}(t,T) \otimes \Phi_{fX}(t,T) \end{pmatrix} \text{ and } \mathcal{M}_{0Y} = \begin{pmatrix} 0_{8\times 1} \\ \Phi_f^2(t,T)\sigma_f^2 \end{pmatrix}. \tag{3.84}$$

With help of simple matrix algebra applied to (3.99), the conditional covariance of Y_t has the (vectorized) form

$$\mathrm{vec}V_Y(t,T) = \int_{s=t}^{T} \mathcal{M}_{1Y}(s,T)\left[\Phi_V(s,T)\otimes\Phi_V(s,T)\right] ds \times \mathrm{vec}\left(V_t\right) \tag{3.85}$$

$$+\int_{s=t}^{T} k\mathcal{M}_{1Y}\mathrm{vec}\left[\bar{\mu}_V(t,s)\right] ds + \int_{s=t}^{T}\mathcal{M}_{0Y}(s,T)ds. \tag{3.86}$$

The unconditional covariance of Y is given as:

$$\lim_{T\to\infty}\mathrm{vec}V_Y(t,T) = \lim_{T\to\infty}\int_{s=t}^{T} k\mathcal{M}_{1Y}(s,T)\mathrm{vec}\left[\bar{\mu}_V(t,s)\right] ds + \int_{s=t}^{T}\mathcal{M}_{0Y}(s,T)ds. \tag{3.87}$$

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

The expressions for the conditional mean (3.75) and covariance (3.86) 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

$$dY_t = \left(\mu_Y + \mathcal{K}_Y Y_t\right) dt + \Sigma_Y\left(V_t\right) dZ_t \tag{3.88}$$

as

$$Y_{t+\Delta t} = \bar{\mu}_{Y,\Delta t} + \Phi_{Y,\Delta t}Y_t + \Sigma_Y\left(V_t\right)\sqrt{\Delta t}\varepsilon_{t+\Delta t}, \tag{3.89}$$

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:

$$\begin{aligned}\mathrm{vec}E\left(YY'\right) &= \left(I - \Phi_{Y,\Delta t}\otimes\Phi_{Y,\Delta t}\right)^{-1}\times \\ &\times\mathrm{vec}\left\{\bar{\mu}_{Y,\Delta t}\bar{\mu}'_{Y,\Delta t} + \bar{\mu}_{Y,\Delta t}E\left(Y'\right)\Phi'_{Y,\Delta t} + \Phi_{Y,\Delta t}E\left(Y\right)\bar{\mu}'_{Y,\Delta t} + E\left[\Sigma_Y\left(V_t\right)\Sigma_Y\left(V_t\right)'\right]\Delta t\right\}\end{aligned} \tag{3.90}$$

$$\mathrm{vec}\left[Var\left(Y\right)\right] = \mathrm{vec}E\left(YY'\right) - \mathrm{vec}E\left(Y\right)\left[\mathrm{vec}E\left(Y\right)\right]'.$$

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.

VIII.D. Model estimation

Discretization and vectorization of the state space

This section collects the details about the vectorization of transition dynamics for Y_t and V_t.

Parameters for discretized transition dynamics of Y_t are given by:

$$\bar{\mu}_{Y,\Delta t} = \left(e^{\mathcal{K}_Y \Delta t} - I\right) \mathcal{K}_Y^{-1} \mu_Y \tag{3.91}$$

$$\Phi_{Y,\Delta t} = e^{\mathcal{K}_Y \Delta t}. \tag{3.92}$$

Parameter matrices $\Phi_{V,\Delta t}$ and $\bar{\mu}_{V,\Delta t}$ for discretized transition dynamics of V_t are given by:

$$\bar{\mu}_{V,\Delta t} = \int_0^{\Delta t} \Phi_{V,s} Q' Q \Phi'_{V,s} ds \tag{3.93}$$

$$\Phi_{V,\Delta t} = e^{M \Delta t}, \tag{3.94}$$

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.

We recast the discretized covariance matrix dynamics V_t in a vector form:

$$\mathrm{vec}(V_{t+\Delta t}) = k\mathrm{vec}(\bar{\mu}_{V,\Delta t}) + (\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathrm{vec}(V_t) + \mathrm{vec}(u^V_{t+\Delta t}). \tag{3.95}$$

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: $\mathcal{E}_n\mathrm{vec}(A) = \mathrm{vech}(A)$, where $\mathrm{vech}(\cdot)$ denotes half-vectorization, (ii) a duplication matrix: $\mathcal{D}_n\mathrm{vech}(A) = \mathrm{vec}(A)$. Using half-vectorization, we define $\bar{V}_t := \mathrm{vech}(V_t) = \mathcal{E}_n\mathrm{vec}(V_t)$, which contains $\bar{n} = n(n+1)/2$ unique elements of V_t:

$$\bar{V}_{t+\Delta t} = k\mathcal{E}_n\mathrm{vec}(\bar{\mu}_{V,\Delta t}) + \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n\bar{V}_t + \mathcal{E}_n\mathrm{vec}(u^V_{t+\Delta t}). \tag{3.96}$$

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

$$E_t\left(S_{t+\Delta t}\right) = \begin{pmatrix} \left(e^{\mathcal{K}_Y \Delta t} - I\right) \mathcal{K}_Y^{-1} \mu_Y + e^{\mathcal{K}_Y \Delta t} Y_t \\ k\mathcal{E}_n\mathrm{vec}(\bar{\mu}_{V,\Delta t}) + \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n\bar{V}_t \end{pmatrix}, \tag{3.97}$$

and the conditional covariance of the form:

$$Cov_t\left(S_{t+\Delta t}\right) = \begin{pmatrix} Cov_t(Y_{t+\Delta t}) & 0_{n\times\bar{n}} \\ 0_{\bar{n}\times n} & Cov_t\left(\bar{V}_{t+\Delta t}\right) \end{pmatrix}. \tag{3.98}$$

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:

$$Cov_t(Y_{t+\Delta t}) = \Sigma_Y(V_t)\Sigma_Y(V_t)'\Delta t \tag{3.99}$$

$$\begin{aligned} Cov_t\left(\bar{V}_{t+\Delta t}\right) &= \mathcal{E}_n Cov_t\left(V_{t+\Delta t}\right)\mathcal{E}_n' \\ &= \mathcal{E}_n\left(I_{n^2} + K_{n,n}\right)\left[\Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}' \otimes \bar{\mu}_{V,\Delta t} + k\left(\bar{\mu}_{V,\Delta t} \otimes \bar{\mu}_{V,\Delta t}\right) + \bar{\mu}_{V,\Delta t} \otimes \Phi_{V,\Delta t}V_t\Phi_{V,\Delta t}'\right]\mathcal{E}_n', \end{aligned} \tag{3.100}$$

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.

Econometric identification

This section details our econometric identification procedure and parameter restrictions. 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.

This section 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:

$$S_{t+\Delta t} = A + BS_t + \varepsilon_{t+\Delta t}, \tag{3.101}$$

and the compound measurement equation is given by:

$$m_t = h(S_t; \Theta) + \vartheta_t. \tag{3.102}$$

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

$$A = \begin{pmatrix} (\Phi_{Y,\Delta t} - I)\,\mathcal{K}_Y^{-1}\mu_Y \\ k \cdot \mathcal{E}_n vec(\bar{\mu}_{V,\Delta t}) \end{pmatrix}. \tag{3.103}$$

B is a block-diagonal matrix of the form:

$$B = \begin{pmatrix} \Phi_{Y,\Delta t} & \mathbf{0}_{n\times\bar{n}} \\ \mathbf{0}_{\bar{n}\times n} & \mathcal{E}_n(\Phi_{V,\Delta t} \otimes \Phi_{V,\Delta t})\mathcal{D}_n \end{pmatrix}. \tag{3.104}$$

The vector shocks is of the form:

$$\varepsilon_{t+\Delta t} = \begin{pmatrix} u^Y_{t+\Delta t} \\ \mathcal{E}_n \text{vec}(u^V_{t+\Delta t}) \end{pmatrix}, \tag{3.105}$$

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

$$Cov_t(\varepsilon_{t+\Delta t}) = \begin{pmatrix} Cov_t(Y_{t+\Delta t}) & \mathbf{0}_{n\times\bar{n}} \\ \mathbf{0}_{\bar{n}\times n} & Cov_t(\bar{V}_{t+\Delta t}) \end{pmatrix}. \tag{3.106}$$

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:

$$h(S_t; \Theta) = \begin{pmatrix} f(S_t; \Theta) \\ g(V_t; \Theta) \end{pmatrix}. \tag{3.107}$$

The vector of measurement errors:

$$\vartheta_t = \begin{pmatrix} \sqrt{R_y} e_t^y \\ \sqrt{R_v} e_t^v \end{pmatrix} \tag{3.108}$$

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

$$Cov(\vartheta_t) = \begin{pmatrix} \sigma_y^2 \mathbf{I}_6 & \mathbf{0}_{6\times 3} \\ \mathbf{0}_{3\times 6} & diag(\sigma_{i,v}^2)_{i=1,2,3} \end{pmatrix}. \tag{3.109}$$

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:

$$\mathcal{S}_0 = \bar{S} \tag{3.110}$$

$$\mathcal{S}_i = \bar{S} + \left(\sqrt{(L_S+\lambda)P_S}\right)_i, i = 1, ..., L_S \tag{3.111}$$

$$\mathcal{S}_i = \bar{S} - \left(\sqrt{(L_S+\lambda)P_S}\right)_{i-L_S}, i = L_S + 1, ..., 2L_S, \tag{3.112}$$

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:

$$\bar{m} \approx \sum_{i=0}^{2L_S} W_i^{\mu} \mathcal{M}_i \tag{3.113}$$

$$P_S \approx \sum_{i=0}^{2L_S} W_i^{\sigma} (\mathcal{M}_i - \bar{m})(\mathcal{M}_i - \bar{m})', \tag{3.114}$$

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

$$W_0^{\mu} = \frac{\lambda}{L_S + \lambda} \tag{3.115}$$

$$W_0^{\sigma} = \frac{\lambda}{L_S + \lambda} + 1 - \alpha^2 + \beta, i = 1, ..., L_S \tag{3.116}$$

$$W_i^{\mu} = W_i^{\sigma} = \frac{\lambda}{2(L_S + \lambda)}, i = L_S + 1, ..., 2L_S. \tag{3.117}$$

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

The UKF Algorithm

1. Initialize at unconditional moments:[26]

$$\hat{S}_0 = \mathbb{E}[S_0] \tag{3.118}$$

$$P_{S0} = \mathbb{E}[(S_0 - \hat{S}_0)(S_0 - \hat{S}_0)'] \tag{3.119}$$

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

2. Compute the sigma points:

$$\mathcal{S}_{k-1} = \begin{bmatrix} \hat{S}_{k-1} & \hat{S}_{k-1} + \sqrt{(L_S + \lambda)P_{S,k-1}} & \hat{S}_{k-1} - \sqrt{(L_S + \lambda)P_{S,k-1}} \end{bmatrix} \tag{3.120}$$

3. Time update:

$$\mathcal{S}^a_{k|k-1} = B(\mathcal{S}_{k-1}) \tag{3.121}$$

$$\hat{S}^-_k = \sum_{i=0}^{2L_S} W^\mu_i \mathcal{S}^a_{k|k-1} \tag{3.122}$$

$$P^-_{Sk} = \sum_{i=0}^{2L_S} W^\sigma_i (\mathcal{S}^a_{ik|k-1} - \hat{S}^-_k)(\mathcal{S}^a_{ik|k-1} - \hat{S}^-_k)' + Cov_t(\varepsilon_{t+\Delta t}) \tag{3.123}$$

4. Augment sigma points:

$$\mathcal{S}_{k|k-1} = \begin{bmatrix} \mathcal{S}^a_{k|k-1} & \mathcal{S}^a_{0k|k-1} + \sqrt{(L_S + \lambda)Cov_t(\varepsilon_{t+\Delta t})} & \mathcal{S}^a_{0k|k-1} - \sqrt{(L_S + \lambda)Cov_t(\varepsilon_{t+\Delta t})} \end{bmatrix} \tag{3.124}$$

$$\mathcal{M}_{k|k-1} = h(\mathcal{S}_{k|k-1}) \tag{3.125}$$

$$\hat{m}^-_k = \sum_{i=1}^{2L_S} W^\sigma_i \mathcal{M}_{i,k|k-1} \tag{3.126}$$

5. Measurement equations update:

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

$$P^-_{m_k} = \sum_{i=0}^{2L_S} W^\sigma_i (\mathcal{M}_{ik|k-1} - \hat{m}^-_k)(\mathcal{M}_{ik|k-1} - \hat{m}^-_k)' + Cov_t(\vartheta_{t+\Delta t}) \tag{3.127}$$

$$P_{S_k m_k} = \sum_{i=0}^{2L_S} W^\sigma_i (\mathcal{S}_{ik|k-1} - \hat{S}^-_k)(\mathcal{M}_{ik|k-1} - \hat{m}^-_k)' \tag{3.128}$$

$$\mathcal{K}_k = P_{S_k m_k} P^{-1}_{m_k} \tag{3.129}$$

$$\hat{F}_k = \hat{F}^-_k + \mathcal{K}_k(m_k - \hat{m}^-_k) \tag{3.130}$$

$$P_k = P^-_k - \mathcal{K}_k P^-_{m_k} \mathcal{K}'_k. \tag{3.131}$$

Pseudo-maximum likelihood estimation

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:

$$l_{t+1}(\Theta) = -\frac{1}{2} \ln |P^-_{m,t+1}| - \frac{1}{2} \left(\hat{m}^-_{t+1} - m_{t+1}\right)' \left(P^-_{m,t+1}\right)^{-1} \left(\hat{m}^-_{t+1} - m_{t+1}\right), \tag{3.132}$$

and obtain parameter estimates by maximizing the criterion:

$$\hat{\Theta} := \arg\min_{\Theta} \mathcal{L}\left(\Theta, \{m_t\}_{t=1}^T\right) \quad \text{with} \quad \mathcal{L}\left(\Theta, \{m_t\}_{t=1}^T\right) = \sum_{t=0}^{T-1} l_{t+1}(\Theta). \tag{3.133}$$

with $T = 1003$ weeks. The initial log-likelihood is evaluated at the unconditional moments of the state vector (see Section VIII.C for the expressions).

Bibliography

ADRIAN, T., AND H. WU (2009): "The Term Structure of Inflation Expectations," Working Paper, Federal Reserve Bank of New York.

ALBUQUERQUE, R., M. EICHENBAUM, AND S. REBELO (2012): "Valuation Risk and Asset Pricing," Working paper, Northwestern University.

ANDERSEN, T. G., AND L. BENZONI (2008): "Realized Volatility," Working Paper, Kellogg School of Management and Federal Reserve Bank of Chicago.

ANDERSEN, T. G., AND L. BENZONI (2010): "Do Bonds Span Volatility Risk in the US Treasury Market? A Specification Test for Affine Term Structure Models," *Journal of Finance*, 65, 603–655.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND C. VEGA (2007): "Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets," *Journal of International Economics*, 73, 251–277.

ANG, A., G. BEKAERT, AND M. WEI (2007a): "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54, 1163–1212.

——— (2007b): "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?," *Journal of Monetary Economics*, 54, 1163–1212.

——— (2007c): "The Term Structure of Real Rates and Expected Inflation," Columbia University, NBER and Federal Reserve Board of Governors.

ANG, A., J. BOIVIN, S. DONG, AND R. LOO-KUNG (2010): "Monetary Policy Shifts and the Term Structure," *Review of Economic Studies*, forthcoming.

ANG, A., M. PIAZZESI, AND M. WEI (2006): "What does the yield curve tell us about GDP growth?," *Journal of Econometrics*, 131, 359–403.

ANGELETOS, G.-M., AND J. LA'O (2012): "Optimal Monetary Policy with Informational Frictions," Working paper, University of Chicago.

Atkeson, A., and P. J. Kehoe (2008): “On the Need for a New Approach to Analyzing Monetary Policy,” *NBER Macroeconomics Annual*, 23.

Attanasio, O., J. Banks, and S. Tanner (2002): “Asset Holding and Consumption Volatility,” *Journal of Political Economy*, 110, 771–792.

Attanasio, O., and H. Low (2004): “Estimating Euler Equations,” *Review of Economic Dynamics*, 7, 406–435.

Audrino, F., and F. Corsi (2007): “Realized Correlation Tick-by-Tick,” Working paper, University of St. Gallen.

——— (2008): “Realized Covariance Tick-by-Tick in Presence of Rounded Time Stamps and General Microstructure Effects,” Working paper, University of St. Gallen.

Azarias, G. (2010): “The Swaption Market,” Morgan Stanley.

Bacchetta, P., E. Mertens, and E. van Wincoop (2009): “Predictability in financial markets: What do survey expectations tell us?,” *Journal of International Money and Finance*, 28, 406–426.

Balke, N., and R. Gordon (1989): “The Estimation of Prewar Gross National Product: Methodology and New Evidence,” *Journal of Political Economy*, 97, 38–92.

Balke, N., and R. J. Gordon (1986): “Appendix B. Historical Data,” in *The American Business Cycle: Continuity and Change*, ed. by R. J. Gordon, pp. 781 – 850. UMI.

Bandi, F., and J. Russell (2008): “Microstructure Noise, Realized Variance, and Optimal Sampling,” *Review of Economic Studies*, 75, 339–369.

Bansal, R., D. Kiku, and A. Yaron (2012): “An Empirical Evaluation of the Long-Run Risk Model for Asset Prices,” *Critical Finance Review*, 1, 183–221.

Bansal, R., and I. Shaliastovich (2010): “A Long-Run Risk Explanation of Predictability Puzzles in Bond and Currency Markets,” Working Paper, Duke University.

Bansal, R., and Y. Yaron (2004): “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance*, 59, p. 1481 – 1509.

Barillas, F., and K. Nimark (2012): “Speculation, Risk Premia and Expectations in the Yield Curve,” Working paper Goizueta Business School and CREI.

Barndorff-Nielsen, O. E., and N. Shephard (2004): “Econometric Analysis of Realised Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics,” *Econometrica*, 72, 885–925.

Barro, R. (2009): “Rare Disasters, Asset Prices, and Welfare Costs,” *American Economic Review*, 99, 243–264.

Barro, R., and C. Sahasakul (1983): "Measuring the Average Marginal Tax Rate from the Individual Income Tax," NBER working paper.

Barsky, R. (1987): "The Fisher Hypothesis and the Forecastability and Persistence of Inflation," *Journal of Monetary Economics*, 19, 3–24.

Bauer, M. D., and G. D. Rudebusch (2011): "The Signaling Channel for Federal Reserve Bond Purchases," Working paper, San Francisco Fed.

Beeler, J., and J. Y. Campbell (2008): "The Long-Run Risk Model and Aggregate Asset Prices: An Empirical Assessment," Working Paper, Harvard University.

Bekaert, G., S. Cho, and A. Moreno (2010): "New Keynesian Macroeconomics and the Term Structure," *Journal of Money, Credit and Banking*, 42, 34–62.

Bekker, P. A., and K. E. Bouwman (2009): "Risk-free Interest Rates Driven by Capital Market Returns," Working Paper, University of Groningen and Erasmus University Rotterdam.

Benati, L., and C. Goodhart (2008): "Investigating time-variation in the marginal predictive power of the yield spread," *Journal of Economic Dynamics and Control*, 32, 1236–1272.

Bernanke, B. (2011): "The Effects of the Great Recession on Central Bank Doctrine and Practice," 56th Economic Conference, Federal Reserve Bank of Boston.

Bernanke, B. S., and A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901–921.

Bilbiie, F. (2008): "Limited asset markets participation, monetary policy and (inverted) aggregate demand logic," *Journal of Economic Theory*, 140, 162–196.

Bilbiie, F., and R. Straub (2012): "Changes in the Output Euler Equation and Asset Markets Participation," *Journal of Economic Dynamics and Control*, forthcoming.

Boivin, J. (2006): "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," *Journal of Money, Credit and Banking*, 35, 1149–1174.

Bru, M.-F. (1991): "Wishart Processes," *Journal of Theoretical Probability*, 4, 725–751.

Buraschi, A., A. Cieslak, and F. Trojani (2008): "Correlation Risk and the Term Structure of Interest Rates," Working Paper, University of Lugano.

Buraschi, A., P. Porchia, and F. Trojani (2010): "Correlation Risk and Optimal Portfolio Choice," *Journal of Finance*, 65, 393–420.

Campbell, J. (2003): "Consumption-based Asset Pricing," in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. Stulz, chap. 13. Elsevier Science.

Campbell, J., and G. Mankiw (1989): "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," NBER working paper.

Campbell, J. Y., and R. J. Shiller (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

Campbell, J. Y., A. Sunderam, and L. M. Viceira (2012): "Inflation Bets or Deflation Hedges? The Changing Risk of Nominal Bonds," Working paper, Harvard Business School.

Carlson, J. (1977): "A Study of Price Forecasts," in *Annals of Economic and Social Measurement*, ed. by NBER, vol. 6, pp. 33–63.

Carr, P., and L. Wu (2007): "Stochastic Skew in Currency Options," *Journal of Financial Economics*, 86, 213–247.

Carroll, C. (2000): "Requiem for the Representative Consumer? Aggregate Implications of Microeconomic Consumption Behavior," *American Economic Review, Paper and Proceedings*, 90, 110–115.

——— (2001): "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)," *Advances in Macroeconomics*, 1.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders (1992): "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47, 1209–1227.

Chari, V., P. Kehoe, and E. McGrattan (2009): "New Keynesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 1, 242–266.

Chetty, R. (2006): "A New Method of Estimating Risk Aversion," *American Economic Review*, 96, 1821–1834.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–475.

Cho, S., and A. Moreno (2011): "The forward method as a solution refinement in rational expectations models," *Journal of Economic Dynamics and Control*, 35, 257–272.

Christoffersen, P., K. Jacobs, L. Karoui, and K. Mimouni (2009): "Non-Linear Filtering in Affine Term Structure Models: Evidence from the Term Structure of Swap Rates," Working paper, McGill University, Goldman Sachs, American University in Dubai.

Cieslak, A., and P. Povala (2011): "Understanding Bond Risk Premia," Working paper, Kellogg School of Management.

——— (2013): "Expecting the Fed," Working paper, Kellogg School of Management and University of Lugano.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, Jaunary 2000, 147–180.

Clark, T., and S. Kozicki (2004): "Estimating Equilibrium Real Interest Rates in Real Time," Working paper, Federal Reserve Bank of Kansas.

Cochrane, J. (1989): "The Sensitivity of Tests of the Intertemporal Consumption to Near-Rational Alternatives," *American Economic Review*, 79, 319–337.

Cochrane, J., and M. Piazzesi (2002): "The Fed and Interest Rates- A High-Frequency Identification," *Americn Economic Review*, 92, 90–95.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

Cogley, T. (2005): "Changing beliefs and the term structure of interest rates: Cross-equation restrictions with drifting parameters," *Review of Economic Dynamics*, 8, 420–451.

Cogley, T., G. Primiceri, and T. Sargent (2010): "Inflation-Gap Persistence in the US," *American Economic Journal: Macroeconomics*, 2, 43–69.

Cogley, T., T. Sargent, and P. Surico (2012a): "Price-Level Uncertainty and Stability in the UK," Working paper, NYU and LBS.

——— (2012b): "The Return of the Gibson Paradox," Working paper, New York University and London Business School.

Cogley, T., and T. J. Sargent (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, 8, 263–302.

Cogley, T., and A. M. Sbordone (2008): "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review*, 98, 2101–2126.

Coibion, O., and Y. Gorodnichenko (2011a): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," Working paper, University of California Berkeley and College of William and Mary.

——— (2011b): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101, 341–370.

Coibion, O., and Y. Gorodnichenko (2012): "What Can Survey Forecasts Tell Us About Informational Rigidities?," *Journal of Political Economy*, 120, 116–159.

Collin-Dufresne, P., and R. S. Goldstein (2002): "Do Bonds Span the Fixed Income Markets? Theory and Evidence for the Unspanned Stochastic Volatility," *Journal of Finance*, 58(4), 1685–1730.

Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones (2009): "Can Interest Rate Volatility Be Extracted from the Cross-Section of Bond Yields?," *Journal of Financial Economics*, 94, 47–66.

Cooper, I., and R. Priestley (2009): "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, 22, 2801–2833.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 373–384.

Cumby, M. C. R., and B. Diba (2007): "Euler equations and money market interest rates: A challenge for monetary policy models," *Journal of Monetary Economics*, 54, 1863–1881.

da Fonseca, J., M. Grasselli, and C. Tebaldi (2007): "Option Pricing when Correlations Are Stochastic: An Analytical Framework," *Review of Derivatives Research*, 10, 151–180.

Dai, Q., and K. Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–1978.

Duarte, J. (2008): "The Causal Effect of Mortgage Refinancing on Interest-Rate Volatility: Empirical Evidence and Theoretical Implications," *Review of Financial Studies*, 21, 1689–1731.

Duffee, G. (2012a): "Bond pricing and the macroeconomy," in *Handbook of the Economics of Finance*, ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Elsevier Science.

——— (2012b): "Forecasting Interest Rates," Working paper, Johns Hopkins University.

Duffee, G. R. (2002): "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405–443.

——— (2009): "Forecasting with the term structure: The role of no-arbitrage restrictions," Working paper, Johns Hopkins University.

——— (2011): "Information in (and Not in) the Term Structure," *Review of Financial Studies*, 24, 2895–2934.

Eichenbaum, M., and J. Fisher (2007): "Estimating the frequency of price re-optimization in Calvo-style models," *Journal of Monetary Economics*, 54, 2032–2047.

Ellison, M., and T. J. Sargent (2010): "A Defence of the FOMC," Working paper, University of Oxford and New York University.

Estrella, A., and G. A. Hardouvelis (1991): "The Term Structure as a Predictor of Real Economic Activity," *Journal of Finance*, 46, 555–576.

Evans, M. D. D., and K. Lewis (1995): "Do Expected Shifts in Inflation Affect Estimates of the Long-Run Fisher Relation?," *Journal of Finance*, 50, 225–253.

Fama, E. (1975): "Short-Term Interest Rates as Predictors of Inflation," *American Economic Review*, 65, 269–282.

Fama, E., and M. Gibbons (1982): "Inflatiom, Rear Return and Capital Investment," *Journal of Monetary Economics*, 9, 297–323.

Fama, E. F. (1990): "Term-structure Forecasts of Interest Rates, Inflation, and Real Returns," *Journal of Monetary Economics*, 25, 59–76.

Fama, E. F., and R. R. Bliss (1987): "The Information in Long-Maturity Forward Rates," *American Economic Review*, 77, 680–692.

Faust, J., and J. Wright (2012): "Forecasting Inflation," Working paper, Johns Hopkins University.

Faust, J., and J. H. Wright (2011): "Forecasting Inflation," Draft for Handbook of Forecasting, Johns Hopkins University.

Ferrero, G., and A. Nobili (2009): "Futures contract rates as monetary policy forecasts," *International Journal of Central Banking*, 5, 109–145.

Fisher, M., and C. Gilles (1996): "Estimating Exponential-Affine Models of the Term Structure," Working Paper, Federal Reserve Bank of Atlanta.

Fisher, M., D. Nychka, and D. Zervos (1994): "Fitting the term structure of interest rates with smoothing splines," Working paper, Federal Reserve and North Carolina State University.

Fleming, M., and K. Garbade (2005): "Explaining settlement fails," *Current Issues in Economics and Finance*, 11, 1–7.

Fleming, M. J. (1997): "The Round-the-Clock Market for U.S. Treasury Securities," *FRBNY Economic Policy Review*.

Fleming, M. J., and B. Mizrach (2009): "The Microstructure of a U.S. Treasury ECN: The BrokerTec Platform," Working paper, Federal Reserve Bank of New York and Rutgers University.

Fontaine, J.-S., and R. Garcia (2011): "Bond Liquidity Premia," *Review of Financial Studies*, forthcoming.

Frankel, J. A., and K. A. Froot (1987): "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations," *American Economic Review*, 77, 133–153.

Friedman, B. M. (1979): "Optimal Expectations and the Extreme Information Assumptions of Rational Expectations Models," *Journal of Monetary Economics*, 5, 23–41.

Froot, K. (1989): "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Finance*, 44, 283–305.

Fuhrer, J., and G. Rudebusch (2004): "Estimating the Euler equation for output," *Journal of Monetary Economics*, 51, 1133–1153.

Fuster, A., D. Laibson, and B. Mendel (2010): "Natural Expectations and Macroeconomic Fluctuations," *Journal of Economic Perspectives*, 24, 67–84.

Gabaix, X. (2012): "Sparse Dynamic Programming and Aggregate Fluctuations," Working paper, NYU Stern.

Gabaix, X., and D. Laibson (2001): "The 6D bias and the equity-premium puzzle," *NBER Macroeconomics Annual*, 16.

Gali, J., D. Lopez-Salido, and J. Valles (2004): "Rule-of-Thumb Consumers and the Design of Interest Rate Rules," *Journal of Money, Credit and Banking*, 36, 739–763.

Gallmeyer, M., B. Hollifield, F. Palomino, and S. Zin (2007): "Arbitrage-Free Bond Pricing with Dynamic Macroeconomic Models," *Federal Reserve Bank of St. Louis Review*, 89, 305–326.

Garbade, K., F. Keane, L. Logan, A. Stokes, and J. Wolgemuth (2010): "The Introduction of the TMPG Fails Charge for U.S. Treasury Securities," *FRBNY Economic Policy Review*, pp. 45–71.

Garcia, R., and P. Perron (1996): "An Analysis of the Real Interest Rate Under Regime Shifts," *Review of Economics and Statistics*, 78, 111–125.

Gourieroux, C. (2006): "Continuous Time Wishart Process for Stochastic Risk," *Econometric Reviews*, 25, 177–217.

Gourieroux, C., J. Jasiak, and R. Sufana (2009a): "The Wishart Autoregressive Process of Multivariate Stochastic Volatility," *Journal of Econometrics*, 150, 167–181.

——— (2009b): "The Wishart Autoregressive Process of Multivariate Stochastic Volatility," *Journal of Econometrics*, 150, 167–181.

Greenwood, R., and A. Shleifer (2013): "Expectations of Returns and Expected Returns," Working paper, Harvard University.

Gruber, J. (2006): "A Tax-based Estimate of Elasticity of Intertemporal Substitution," NBER working paper.

Gurkaynak, R., B. Sack, and E. T. Swanson (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1, 55–93.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2006): "The U.S. Treasury Yield Curve: 1961 to the Present," Working paper, Federal Reserve Board.

Guvenen, F. (2006): "Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective," *Journal of Monetary Economics*, 53, 1451–1472.

Hall, R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.

Hansen, L., and K. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

——— (1984): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 52, 267–268.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116, 260–302.

Hansen, L. P., and K. J. Singleton (1996): "Efficient Estimation of Linear Asset-Pricing Models with Moving Average Errors," *Journal of Business and Economic Statistics*, 14, 53–68.

Harvey, C. R. (1988): "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 22, 305–333.

——— (1989): "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal*, September–October, 38–45.

Haubrich, J., G. Pennacchi, and P. Ritchken (2011): "Estimating Real and Nominal Term Structures using Treasury Yields, Inflation, Inflation Forecasts, and Inflation Swap Rates," Working Paper, Federal Reserve Bank of Cleveland.

Hautsch, N., and Y. Ou (2008): "Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia," Working Paper, Humboldt Univeristy Berlin.

Hayashi, T., and N. Yoshida (2005): "On covariance estimation of non-synchronously observed diffusion processes," *Bernoulli*, 11, 359–379.

Heidari, M., and L. Wu (2003): "Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?," *Journal of Fixed Income*, 13, 75–86.

Hu, X., J. Pan, and J. Wang (2013): "Noise as Information for Illiquidity," *Journal of Finance*, forthcoming.

Jacobs, K., and L. Karoui (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," Working Paper, Universite Pierre et Marie Curie.

Jagannathan, R., and Y. Wang (2007): "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns," *Journal of Finance*, 62, 1623–1661.

Jones, C. M., O. Lamont, and R. L. Lumsdaine (1998): "Macroeconomic News and Bond Market Volatility," *Journal of Financial Economics*, 47, 315–337.

Joslin, S. (2010): "Pricing and Hedging Volatility Risk in Fixed Income Markets," Working paper, USC Marshall School of Business.

Joslin, S., A. Le, and K. Singleton (2013): "Gaussian Macro-Finance Term Structure Models with Lags," Working paper, Stanford, USC and UNC.

Joslin, S., M. Priebsch, and K. Singleton (2010): "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," Working paper, Stanford GSB, University of Southern California.

JP Morgan (2011): "The Domino Effect of a US Treasury Technical Default," *US Fixed Income Strategy*, April 19, 2011.

Julier, S., and J. Uhlmann (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*.

Justiniano, A., and G. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98, 604–641.

Kim, D. H., and K. J. Singleton (2011): "Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields," Working paper, Stanford GSB, NBER and Yonsei University.

Kohn, D. L. (2005): "Monetary Policy Perspectives on Risk Premiums in Financial Markets," Remarks at the Financial Market Risk Premiums Conference, Washington D.C.

Kozicki, S., and P. Tinsley (2006): "Survey-Based Estimates of the Term Structure of Expected U.S. Inflation," Working paper, Bank of Canada.

Kuttner, K. (2003): "Dating Changes in the Federal Funds Rate, 1989-92," Note, NY Fed.

Kuttner, K. N. (2001): "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market," *Journal of Monetary Economics*, 47, 523–544, Working paper.

Laubach, T., and J. C. Williams (2003): "Measuring the Natural Rate of Interest," *Review of Economics and Statistics*, 85, 1063–1070.

Lebergott, S. (1996): *Consumer Expenditures: New Measures and Old Motives*. Princeton University Press.

Li, H., and F. Zhao (2006): "Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives," *Journal of Finance*, 61, 341–378.

Lucas, R. (1980): "Two Illustrations of the Quantity Theory of Money," *American Economic Review*, 70, 1005–1014.

Ludvigson, S. C., and S. Ng (2009): "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, forthcoming.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2012): "The Wealth-Consumption Ratio," Working paper, NYU, UCLA and MIT.

Magnus, J. R., and H. Neudecker (1979): "The Commutation Matrix: Some Properties and Applications," *Annals of Statistics*, 7, 381–394.

Malloy, C., T. Moskowitz, and A. Vissing-Jorgensen (2009): "Long-Run Stockholder Consumption Risk and Asset Returns," *Journal of Finance*, 64, 2427–2479.

Mankiw, G. (2000): "Savers-Spenders Theory of Fiscal Policy," *American Economic Review*, 90, 120–125.

Mankiw, G., and R. Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117, 1295–1328.

Mankiw, N. G., and R. Reis (2011): "Imperfect Information and Aggregate Supply," *Handbook of Monetary Economics*, 3A, 183–229.

Miron, J. (1986): "Financial Panics, the Seasonality of the Nominal Interest Rate and the Founding of the Fed," *American Economic Review*, 76, 125–140.

Mishkin, F. (1980): "Is the Preferred-Habitat Model of the Term Structure Inconsistent with Financial Market Efficiency?," *Journal of Political Economy*, 88, 406–411.

——— (1981): "The Real Rate of Interest: An Empirical Investigation," *Carnegie-Rochester Conference Series on Public Policy*, 15, 151–200.

Mishkin, F. S. (1990): "The Information in the Longer Maturity Term Structure About Future Inflation," *Quarterly Journal of Economics*, 105, 815–828.

Mizrach, B., and C. J. Neely (2006): "The Transition to Electronic Communications Networks in the Secondary Treasury Market," *Federal Reserve Bank of St. Louis Review*.

Modigliani, F., and R. J. Shiller (1973): "Inflation, Rational Expectations and the Term Structure of Interest Rates," *Economica*, 40, 12–43.

Modigliani, F., and R. Sutch (1966): "Innovations in Interest Rate Policy," *American Economic Review*, 56, 178–197.

——— (1967): "Debt Management and the Term Structure of Interest Rates: An Empirical Analysis of Recent Experience," *Journal of Political Economy*, 75, 569–589.

Nagel, S. (2012): "Macroeconomic Experiences and Expectations: A Perspective on the Great Recession," Working paper, Stanford University.

Neely, C. J., and D. E. Rapach (2008): "Real Interest Rate Persistence: Evidence and Implications," *Federal Reserve Bank of St. Louis Review*.

Neely, C. J., A. Roy, and C. H. Whiteman (2001): "Risk Aversion versus Intertemporal Substitution: A Case Study of Identification Failure in the Intertemporal Consumption Capital Asset Pricing Model," *Journal of Business and Economic Statistics*, 19, 395–403.

Ogaki, M., and C. M. Reinhart (1998): "Measuring Intertemporal Substitution: The Role of Durable Goods," *Journal of Political Economy*, 106, 1078–1098.

Orphanides, A., and J. C. Williams (2005): "Inflation Scares and Forecast-Based Monetary Policy," *Review of Economic Dynamics*, 8, 498–527.

Pastor, L., and R. Stambaugh (2003): "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111, 642–685.

Phoa, W. (1997): "Can You Derive Market Volatility Forecast from the Observed Yield Curve Convexity Bias?," *Journal of Fixed Income*, pp. 43–53.

Piazzesi, M., and M. Schneider (2011): "Trend and Cycle in Bond Premia," Working Paper, Stanford University.

Piazzesi, M., and E. T. Swanson (2008): "Futures Prices as Risk-adjusted Forecasts of Monetary Policy," *Journal of Monetary Economics*, 55, 677–691.

Plosser, C., and G. Rouwenhorst (1994): "International term structures and real economic growth," *Journal of Monetary Policy*, 33, 133–155.

Poterba, J., and A. Samwick (1995): "Stock Ownership Patterns, Stock Market Fluctuations, and Consumption," *Brookings Papers on Economic Activity*, 2, 295–372.

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

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

Reynard, S., and A. Schabert (2012): "Monetary Policy, Interest Rates, and Liquidity Premia," Working paper, Swiss National Bank.

Rhode, P. (2002): "Gallman's Annual Output Series for the United States, 1834-1909," NBER working paper.

Romer, C. (1986): "Is the Stabilization of the Postwar Economy a Figment of the Data?," *American Economic Review*, 76, 314–334.

Romer, C. D., and D. H. Romer (2000): "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90, 429–457.

——— (2008): "The FOMC versus the Staff: Where Can Monetary Policymakers Add Value?," *American Economic Review: Papers and Proceedings*, 98, 230–235.

Rose, A. (1988): "Is the Real Interest Rate Stable?," *Journal of Finance*, 43, 1095–1112.

Rotemberg, J., and M. Woodford (1999): "The Cyclical Behavior of Prices and Costs," in *Handbook of Macroeconomics*, ed. by J. Taylor, and M. Woodford, pp. 1052–1131. Elsevier.

Rudebusch, G. (2002): "Term structure evidence on interest rate smoothing and monetary policy inertia," *Journal of Monetary Economics*, 49, 1161–1187.

Rudebusch, G., and E. Swanson (2012): "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *American Economic Journal: Macroeconomics*, 4, 1–43.

Rudebusch, G. D. (1998): "Do Measures of Monetary Policy in a VAR Make Sense?," *International Economic Review*, 39, 907–931.

Sack, B. (2004): "Extracting the Expected Path of Monetary Policy from Futures Rates," *Journal of Futures Markets*, 24, 733–754.

Sargent, T. (1972): "Rational Expectations and the Term Structure of Interest Rates," *Journal of Money, Credit and Banking*, 4, 74–97.

Sellon, G. (2008): "Monetary Policy Transparency and Private Sector Forecasts: Evidence from Survey Data," *Economic review*, Third quarter 2008, 7–34.

SIFMA (2010): "Research Quarterly, Third Quarter 2010," http://www.sifma.org/research/item.aspx?id=2234.

Singleton, K. J. (2012): "Investor Flows and the 2008 Boom/Bust in Oil Prices," Working paper, Stanford University.

Stock, J., and M. Yogo (2002): "Testing for Weak Instruments in IV Linear Regression," NBER working paper.

Stock, J. H., and M. W. Watson (1998): "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series," NBER working paper.

——— (1999): "Forecasting Inflation," *Journal of Monetary Economics*, 44, 293–335.

——— (2007): "Why Has U.S. Inflation Become Harder to Forecast?," *Journal of Money, Credit, and Banking*, 39, 3–33.

Swanson, E. T., and J. C. Williams (2012): "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates," Working paper, San Francisco Fed.

Trolle, A. B., and E. S. Schwartz (2009): "A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives," *Review of Financial Studies*, 22, 2007–2057.

Trolle, A. B., and E. S. Schwartz (2011): "The Swaption Cube," Working paper, EPFL Lausanne and UCLA Anderson School of Management.

Van Loan, C. F. (1978): "Computing Integrals Involving Matrix Exponential," *IEEE Transactions on Automatic Control*, 23, 395–404.

Vayanos, D., and P.-O. Weil (2008): "A Search-Based Theory of the On-the-Run Phenomenon," *Journal of Finance*, 63, 1361–1398.

Vissing-Jorgensen, A. (2002): "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 110, 825–853.

Vissing-Jorgensen, A. (2004): "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," *NBER Macroeconomics Annual 2003*, 18.

Vissing-Jorgensen, A., and O. Attanasio (2003): "Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion," *American Economic Review, papers and proceedings*, 93, 383–391.

Wachter, J. (2006): "A consumption-based model of the term structure of interest rates," *Journal of Financial Economics*, 79, 365–399.

Waggoner, D. F. (1997): "Spline Methods for Extracting Interest Rate Curves from Coupon Bond Prices," Working paper, Federal Reserve Bank of Atlanta.

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

Weil, P. (1989): "The equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

Whiteman, C. (1984): "Lucas on the Quantity Theory: Hypothesis Testing without Theory," *American Economic Review*, 74, 742–749.

Wolff, E. (2004): "Changes in Household Wealth in the 1980s and 1990s in the U.S.," in *International Perspectives on Household Wealth*, ed. by E. Wolff. Elgar Publishing Ltd.

Woodford, M. (2003a): *Imperfect Common Knowledge and the Effects of Monetary Policy*chap. Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Princeton University Press.

——— (2003b): *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

——— (2010): "Robustly Optimal Monetary Policy with Near-Rational Expectations," *American Economic Review*, 100, 274–303.

——— (2012): "Macroeconomic Analysis without the Rational Expectations Hypothesis," *Annual Review of Economics, forthcoming*, Ann.

Wright, J. (2011): "Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior," Working Paper, Johns Hopkins University.

Yogo, M. (2004): "Estimating the Elasticity of Intertemporal Substitution when Instruments are Weak," *Review of Economics and Statistics*, 86, 797–810.

Zhang, L., P. A. Mykland, and Y. Ait-Sahalia (2005): "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411.