Predictability From a Macroeconomic Term Structure Model

Satadru Hore

Department of Supervision, Regulation and Credit

The Federal Reserve Bank of Boston[1]

[1]Email:satadru.hore@bos.frb.org. Phone: (617)-973-3363.

Abstract

Mixing information from macroeconomic fundamentals and yields helps to understand bond return predictability. We derive the term-structure of interest rates from a structural setting with CRRA preferences and flexible dynamics of macroeconomic growth rates that puts formal restrictions between factor dynamics and risk premia. By letting conditional volatilities of the growth rates fluctuate through time, we generate a quadratic term-structure model along with time-variation in bond risk-premia that is determined endogenously. Estimating the parameters from yield curve and the macroeconomic fundamentals jointly, we conclude that by effectively using prior information to mix bond yields and macroeconomic variables, we can predict bond returns better than from either matching time-series of yields or macroeconomic fundamentals. The posterior distribution of R-squares obtained from this partial equilibrium setting can be as high as 8-10% on bond returns of maturity 2-5 years. Furthermore, bond risk-premia is small and largely negative in the post-Volcker regime.

Keywords: Equilibrium Yield Curve, Bayesian MCMC, Filtering, State-space models, Bond return predictability. JEL Classification Code: C11, G12.

Introduction

The term-structure of interest rates is a curiousity of both financial economists and macroeconomists. The former studies the term-structure under powerful no-arbitrage models (see for example Duffie and Kan (1996) and Dai and Singleton (2000)), but the precise relationship of the short-rate with macroeconomic variables is left unexplored. Macroeconomists are interested in knowing how the Fed conducts its monetary policy and what information the yield curve conveys in determining the Fed's decision. Of late, a literature has emerged that tries to put the two together - by constraining the short-rate to economic variables and then imposing the effect of no-arbitrage on this short-rate to determine the yield curve. This paper imposes one more structural restriction by introducing preferences that imposes constraints between factor dynamics and the underlying risk-premia in the economy. Empirically, we use priors to mix information from macroeconomic variables and yield curve, and that helps us in predicting bond return. From partial equilibrium dynamics, we get posterior distribution of R^2's that rise as high as 8-10% on bond returns of maturity 2-5 years. This is similar to the R^2's arising from macroeconomic yield curve models like Ang, Dong and Piazzesi (2007) (henceforth ADP (2007)) and Campbell, Sunderam and Viceira (2008) (henceforth CSV (2008)).

The macroeconomic term-structure literature has many flavors. Some start with a structural model of the economy that gives rise to a version of Taylor rule, and then impose no-arbitrage by exogenously specifying the dynamics of market prices of risk. For example, Gallmeyer, Hollifield and Zin (2005), Hordahl, Tristani and Vestin (2006), etc. fall into this category. Another strand of the literature starting with Sargent (1979) has tried to estimate VAR systems of yields under the assumption of Expectations Hypothesis. Following Campbell and Shiller (1991)'s evidence on the failure of the Expectations Hypothesis, Ang and Piazzesi (2003) start with a VAR specification of underlying macroeconomic dynamics, posit the short-rate to be affine in these variables and build a no-arbitrage model of the yield curve by exogenously specifying the market prices of risk to address the failure of the Expectations Hypothesis. They uncover that imposing no-arbitrage restrictions improve out-of-sample forecasts over VAR predictions. Recent work by ADP (2007) show that the VAR-affine dynamics is flexible enough to accomodate many different formulations of Taylor rules, and they provide a unifying framework to study different flavors of such interest rate rules. Here, as well, the authors exogenously specify the prices of risk.

In this paper, we propose a dynamic model of macroeconomic fundamentals, specify preferences and impose no-aritrage to determine the yield curve. We assume that agents have CRRA preferences and we take as given the joint dynamics of consumption and CPI growth. Expected consumption growth and expected inflation follow a VAR system with dynamic feedback effects, and the conditional volatilities of the growth rates are further time-varying. We show that even though the nominal pricing kernel has iid shocks, the presence of time-varying conditional volatilities of the growth rates produce time-varying risk premia. In other words, even though the market prices of risk are constant, the quantity of risk is time-varying. Ultimately, this setting produces an yield curve that is linear in the growth rates but quadratic in the conditional volatilities of the growth rates making it a quadratic term-structure model in the tradition of Longstaff (1989), Constantinides (1992), Ahn, Dittmar and Gallant (2002), and others. Recently, CSV (2008) also explores a quadratic term structure model with reduced form market prices of risk where the covariation between inflation and the real pricing kernel can switch-signs, is non-linear in the volatilities and can generate a quadratic yield curve. The assumption of CRRA preferences greatly facilitates the derivation of a closed-form solution of the yield curve which we take directly to the data. There is a long literature dating back to Mehra and Prescott (1985) and Weil (1989) that shows the inability of time-separable CRRA preferences to match observed risk-premium from stocks and risk-free rates jointly. The failure occurs in trying to determine aversion from risk and preference for certainty equivalence of consumption jointly through the curvature parameter of time-separable CRRA preferences. In our context, we do not have this simultaneous problem. If we were to explain equity premium (aversion from an unknown stock price in the future) and bond yields (shifting non-stochastic consumption across time) jointly, then we would need some new machinery like the utility functions due to Epstein and Zin, or Duffie and Epstein. However, since our focus is on determining bond yields only, we are unlikely to face the conundrums that gave rise to the equity-premium/risk-free rate puzzle.

The main contribution of this paper is to estimate parameters and state variables from the above equilibrium quadratic yield curve with underlying consumption and CPI growth. We estimate our parameters and filter state variables jointly from the yield curve and macroeconomic variables by preserving discipline on the coefficients in the yield curve that is imposed by underlying dynamics, preferences and the no-arbitrage restriction. Our main finding is that combining macroeconomic variables with yield curve dynamics generates risk-premia that can explain bond returns much better than if we extracted premia from either source. This idea of estimating parameters and state variables by mixing information from both macro fundamentals and the yield curve that are determined jointly is at the heart of any economic term-structure model. Like ADP (2007), we take a Bayesian approach, and use priors informatively in order to mix information from the macroeconomic fundamentals and the yield curve. By changing priors we can accomodate different exposures to each series. For example, we can use priors to put a heavy emphasis on the yield curve. We can fit the yield curve very tightly like a reduced form model, but the underlying latent variables that are filtered in that process bear little resemblance to the underlying economic structure. Similarly, we can generate parameters and latent shocks to fit the time-series of the macroeconomic fundamentals, but they cannot generate the time-series fluctuations of the yield curve. However, using a moderate prior that pulls in information from both the macroeconomic variables and the yield curve, we generate parameters and latent variables that can not only match the time-series of yields and macroeconomic fundamentals reasonably well, but also generate risk-premia that can explain bond returns with economically significant R^2's. The setting that fits the yield curve too strongly generates risk-premia that is too high and *negatively* correlated with actual bond returns, whereas the setting that fits macroeconomic fundamentals closely simply cannot generate any statistically significant risk-premia. Thus, exposure to both data sources through the help of priors helps us to address key model implications that are implied by a dynamic term-structure model like predictability. The filtering of state variables is a particularly challenging problem in our setting because of the non-linearity in the yield curve and the underlying macroeconomic dynamics. Hore, Lopes and McCulloch (HLM) (2009)[2] develop a non-linear filtering algorithm by generalizing the logic behind Forward Filtering and Backward Sampling (FFBS) developed in Carter and Kohn (1994) and Fruhwirth-Schnatter (1994). We use a particular version of the HLM (2009) methodology adopted to fit the non-linearities of the quadratic term-structure generated by underlying non-linearities in the growth rate dynamics.

Econometrically, we are similar to ADP (2007) because we estimate both the yield curve and the macroeconomic fundamentals jointly, but our filtering problems are very different. Their filtering task is strictly linear and they use standard Bayesian filtering mechanisms to filter and smooth an univariate latent variable. We have two filtering problems - one linear and the other non-linear. First of all, our model implies a forward-looking Taylor rule similar to (Clarida, Gali and Gertler, 2000) where the nominal rate is a linear function of expected inflation and expected consumption growth - none of which are observable in our model. Furthermore, the conditional volatilities of each of these latent variables are also unobservable and they enter non-linearly in the yield curve as well as growth dynamics. In our filtering exercise, we filter all four of these latent variables. Whereas filtering the growth rates is a linear filtering problem for which we use the standard FFBS of Carter and Kohn (1994), the non-linear filtering problem is executed using the generalized FFBS of HLM (2009).

This paper is organized as follows - section 1 describes the model and sets up the state-

[2]available at http://www.rob-mcculloch.org/some_papers_and_talks/papers/working/hore_lopes_mcculloch.pdf

space setting, section 2 discusses the prior settings we analyze and briefly discuss the empirical methodology, section 3 discusses the findings and section 4 concludes.

1 The Model

Assume real consumption, c_t, and CPI, q_t, follow the joint process

$$\frac{dc_t}{c_t} = \mu_t^c\, dt + \sigma_c\, dW_c \tag{1}$$

$$\frac{dq_t}{q_t} = \mu_t^q\, dt + \sigma_q\, dW_q \tag{2}$$

where $corr(dW_c, dW_q) = \rho_{cq} dt$. Expected consumption growth and expected inflation follow predictable processes with time-varying conditional volatilities. The dynamics of the expected growth rates and their conditional volatilities are

$$d\mu_t^c = (\alpha_0^c - \alpha_1^c \mu_t^c - \alpha_2^c \mu_t^q)\, dt + \vartheta_t^c\, dB_c \tag{3}$$

$$d\mu_t^q = (\alpha_0^q - \alpha_1^q \mu_t^c - \alpha_2^q \mu_t^q)\, dt + \vartheta_t^q\, dB_q \tag{4}$$

and

$$d\vartheta_t^c = (\beta_0^c - \beta_1^c \vartheta_t^c)\, dt + \epsilon_c\, dZ_c \tag{5}$$

$$d\vartheta_t^q = (\beta_0^q - \beta_1^q \vartheta_t^q)\, dt + \epsilon_q\, dZ_q \tag{6}$$

where $corr(dW_c, dB_c) = \rho_1$ and $corr(dW_q, dB_q) = \rho_2$. Expected growth rates ((3)-(4)) follow a joint process where expected inflation directly enters into expected consumption growth, and vice-versa. Their innovations (dB_c and dB_q) are independent, but conditional volatilities of the joint processes are time-varying (ϑ_t^c and ϑ_t^q) and follow independent Gaussian processes ((5)-(6)). All the time-varying latent growth rates and their conditional volatilities are known to the investor. For simplicity, the covariance between other conditional volatilities have been shut down to focus on premia generated by conditional volatilities to expected consumption growth and expected inflation.

The interaction between the growth rates is similar to the VAR systems in the reduced form term-structure literature (Dai and Singleton (2000,2002), Duffee (2000), Duffie and Singleton (1997), among others) where the goal is to come up with a flexible factor structure that would still produce an affine, or a generally tractable solution. In this context, it allows us to incorporate valuable macroeconomic phenomenon into the yield curve under a structural setting. For example, (3) suggests that expected inflation will affect consumption growth and will thus affect real interest rates. Thus, $\alpha_2^c \neq 0$ is a violation of the Fisher hypothesis which posits independence between real rates and expected inflation. In discrete-time, ((3)-(4)) reduces to a system of VARs that allow for dynamic feedback between the growth rates of the macroeconomic fundamentals. The key difference between the VAR term-structure literature, e.g. ADP (2007), Ang and Piazzesi (2003), Diebold, Rudebusch and Aruoba (2006) and this paper is that *all* the time-series variables are latent. They are related to underlying macroeconomic variables, but ultimately they manifest themselves in the yield curve, which is determined endogenously.

The innovations in the growth rates are time-varying and follow independent univariate Vasicek processes ((5)-(6)). The motivation for stochastic volatilities in the system of VARs that describe evolution of macroeconomic quantities goes back to Bernanke and Mihov (1998a, 1998b), and Stock's (2001) comments on Cogley and Sargent (2001). Stock (2001) points out that Cogley and Sargent's (2001) measure of regime shifts in inflation could be due to time-varying conditional volatilities in their underlying dynamical system. The setting here incorporates the view of Sims (2001) that the conditional volatilities to the system of VAR's are time-varying. Notice that since the conditional volatilities follow arithmetic processes, their sign can be positive or negative. Identification is provided by the endogenously determined yield curve as these conditional volatilities determine the sign of time variation in risk-premia.

Let the investor be endowed with CRRA preferences of the form $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$, where γ is relative risk-aversion. Furthermore, let the real pricing kernel be $M_t = e^{-\phi t} u'(c_t)$, whereas the nominal pricing kernel be defined by $M_t^n = \frac{M_t}{q_t}$, where ϕ is the time-preference parameter. Given the utility function, the pricing kernel, the consumption (1) and price processes (2), the nominal and real short rate process for the investor follows:

Proposition 1.1 *The nominal short rate process is*

$$r_t^n = \gamma \mu_t^c + \mu_t^q + k \tag{7}$$

where $k = -[\frac{1}{2}\gamma(1+\gamma)\sigma_c^2 + \sigma_q^2 + \rho_{cq}\gamma\sigma_c\sigma_q - \phi]$. *The corresponding real rate is*

$$r_t = \gamma \mu_t^c + \phi - \frac{1}{2}\gamma(1+\gamma)\sigma_c^2 \tag{8}$$

The nominal (M_t^n) *and real* (M_t) *pricing kernels are given by*

$$\begin{aligned} \frac{dM_t^n}{M_t^n} &= -r_t^n dt - \gamma\sigma_c dW_c - \sigma_q dW_q \\ \frac{dM}{M} &= -r_t dt - \gamma\sigma_c dW_c \end{aligned}$$

Proof: In Appendix.

The interest-rate model is similar to a forward looking Taylor rule consistent with Clarida and Gertler (1998), Clarida, Gali and Gertler (2000) and ADP (2007). The nominal short-rate in (7) can be written as

$$r_t^n = k + \gamma E_t\left[\frac{dc_t}{c_t}\right] + E_t\left[\frac{dq_t}{q_t}\right]$$

ADP(2007) identify that there are many Taylor rule specifications that map into the same affine term structure dynamics in their model. The distinction here in this structural setting is that the underlying fundamentals create a non-linear term-structure model due to the stochastic innovations in the growth rates.

In differential form, the real short-rate can be written as

$$dr_t = (a_t - \alpha_1^c r_t)\, dt + \gamma \vartheta_t^c \, dB_c(t) \tag{9}$$

where $a_t = \gamma\alpha_0^c + \alpha_1^c(\phi - \frac{1}{2}\gamma(1+\gamma)\sigma_c^2) - \gamma\alpha_2^c\mu_t^q$. The effect of expected inflation on the real rate is visible through the long-run mean, a_t, of the real short-rate. If Fisher hypothesis were true and expected inflation and the real rate were independent, then α_2^c would be zero and the real short-rate process will be independent of expected inflation. But, a failure of Fisher hypothesis would imply that the date t long-run mean of the real rate is determined by expected inflation. Whether or not the long-run mean is increasing or decreasing will depend on the sign of α_2^c. If $\alpha_2^c < 0$ ($\alpha_2^c > 0$), then a higher inflation expectation would increase (decrease) the long-run mean of the real rate which would also be inconsistent (consistent) with the Mundell-Tobin effect. Similarly, the nominal short-rate can be written as

$$dr_t^n = (a_t^n - \alpha_1^c r_t^n)\, dt + \gamma\vartheta_t^c \, dB_c(t) + \vartheta_t^q \, dB_q(t) \tag{10}$$

where

$$a_t^n = (\gamma\alpha_0^c + \alpha_0^q - \alpha_1^c k) - (\gamma\alpha_2^c + \alpha_2^q - \alpha_1^c)\mu_t^q - \frac{\alpha_1^q(r_t - \phi + \frac{1}{2}\gamma(1+\gamma)\sigma_c^2)}{\gamma}$$

The nominal interest rate process written in differential form shows it is a four-factor model. Two of them, conditional volatility of expected inflation and expected consumption growth, determine the conditional volatility of the nominal short-rate. The other two, expected consumption growth and expected inflation, drive the contamporenous nominal short-rate.

A nominal bond pays \$1 at time T, whereas the real bond pays $\$q_T/q_t$, which is by construction 1-real dollar. Thus, following Yared and Veronesi (2000) the price of a nominal P^n or real P^r bond at time t with maturity $\tau = T - t$ in this setting is given by the standard asset-pricing relationship

$$E_t[dP^n] = r_t^n P^n dt - E_t\left[\frac{dM^n}{M^n} dP^n\right] \quad (11)$$

$$E_t[dP^r] = r_t P^r dt - E_t\left[\frac{dM}{M} dP^r\right] \quad (12)$$

(11) suggests that the change in nominal bond price is due to nominal interest rate and a risk-premia term. The real bond price in (12) offers compensation for consumption risk, but since it shields the buyer from inflation it doesn't offer any premia for exposure to inflation risk. The risk-premia term for either bond is there to accomodate the failure of expectations hypothesis by introducing a compensation for exposure to aggregate risk expressed in the real and nominal pricing kernels in (1.1). The failure of expectations hypothesis is widely known in the literature (e.g. Campbell and Shiller (1991)) and the current macroeconomic term-structure models (ADP (2007), Bikbov and Chernov(2006), Ang and Piazzesi (2003)) address risk-premia by exogenously specifying market prices of risk that are affine functions of the underlying state variables that simultaneously also determine the nominal rate. We differ here in two ways. First, whereas expected consumption growth and expected inflation determine the level of the nominal or real rate, the conditional volatilities of these growth rates determine risk-premia. Furthermore, the setting here puts formal restrictions between risk-premia and underlying dynamics. The market prices of risk in this economy are constant, but the quantity of risk is time-varying. Due to CRRA preferences and the dynamics of consumption and price growth in ((1)-(2)), the market prices of risk are simply $\gamma\sigma_c$ and σ_q. However, time-varying conditional volatilities of expected consumption growth and expected inflation are responsible for generating the time-varying risk-premia.

The derivation of prices is now straight-forward.

Proposition 1.2 *The equilibrium no-arbitrage price of a nominal bond at time t that matures at $\tau = T - t$ with consumption and CPI processes given by (1)-(2), the latent growth rates (3)-(4), volatility given by (5)-(6) and pricing kernel given in (1.1)*

$$\begin{aligned} P_t^n(\tau) &= \exp(A(\tau,\Theta) + B(\tau,\Theta)\mu_t^c + C(\tau,\Theta)\mu_t^q + D(\tau,\Theta)\vartheta_t^c + E(\tau,\Theta)\vartheta_t^q + \\ &\quad F(\tau,\Theta)\vartheta_t^{c^2} + G(\tau,\Theta)\vartheta_t^{q^2}) \end{aligned} \tag{13}$$

where Θ is the entire parameter space governing ((1)-(6)), risk aversion γ, time-preference parameter ϕ and $A \cdots G$ constitute a system of ordinary differential equations with initial condition $A(0) = \cdots = G(0) = 0$ that ensures a payoff of $P_T^n(0) = 1$.

Price of a real bond $\hat{P}$ with similar specification is also exponentially affine with a different set of ODE's $\hat{A} \cdots \hat{G}$.

Define yield of a nominal bond at time t with maturity τ to be $y_t^n(\tau) = -\frac{\ln P_t^n(\tau)}{\tau}$. In this case, it follows from (13) that the nominal yield curve can be written as

$$\begin{aligned} y_t^n(\tau) &= -\frac{1}{\tau}(A(\tau,\Theta) + B(\tau,\Theta)\mu_t^c + C(\tau,\Theta)\mu_t^q + D(\tau,\Theta)\vartheta_t^c + E(\tau,\Theta)\vartheta_t^q + \\ &\quad F(\tau,\Theta)\vartheta_t^{c^2} + G(\tau,\Theta)\vartheta_t^{q^2}) \end{aligned} \tag{14}$$

The yield curve is linear in the growth rates and quadratic in the conditional volatilities of the growth rates. Thus far, with the exception of CSV (2008), the empirical macroeconomic term-structure literature has focussed on macroeconomic dynamics that generate an affine model of the yield curve. The distinction with CSV (2008) is that this is a partial equilibrium dynamics, whereas they rely on exogenous specification of risk prices like the affine macro term-structure literature.

Quadratic latent factor pricing models date back to Longstaff (1989), Beaglehole and Tenney (1992), Constantinides (1992) and Ahn, Dittmar and Gllant (2002). Constantinides'

(1992) model posits a pricing kernel that is quadratic in the underlying state-variables which follow conditionally Gaussian processes as in here. The short-rate is itself quadratic in the underlying state-variables, which gives rise to a quadratic term structure. Similarly, Ahn, et. al. (2002) also start by positing the nominal short-rate to be quadratic in the underlying state-variables and empirically test the strength of quadratic term structure models and find that they outperform affine models in terms of their goodness of fit. However, none of these quadratic term-structure models embed themselves into a macroeconomic setting and estimate the full model along with underlying macroeconomic dynamics. In particular, inference of state variables in this literature is severely constrained by non-linearties. In fact, Ahn, et. al. (2002) circumvent the filtering problem by using efficient method of moments proposed by Gallant and Tauchen (1998). This paper develops a novel method by generalizing the theory behind Kalman filters to estimate the non-linear state variables. CSV (2008) also faces a non-linear filtering problem which they estimate by using the unscented Kalman filter of Julier and Uhlmann (1997). The method used here is fairly general and can be applied to a host of filtering problems with non-linearity in transition density of states and/or observation equations.

The expected return from a nominal and a real bond can be obtained from the pricing equation (13).

$$E_t\left[\frac{dP^n}{P^n}\right] = (r_t^n + \rho_1\gamma\sigma_c B(\tau)\vartheta_t^c + \rho_2\sigma_q C(\tau)\vartheta_t^q)\,dt \tag{15}$$

$$E_t\left[\frac{dP^r}{P^r}\right] = (r_t + \rho_1\gamma\sigma_c \hat{B}(\tau)\vartheta_t^c)\,dt \tag{16}$$

Expected return on a nominal bond can be decomposed into three parts - the first component is due to the prevailing nominal rate, and the other two are premia for holding a nominal bond that is exposed to time-varying volatilities in expected consumption growth and expected inflation. The predictability relationship above tells us that in response to a changes in conditional volatilities to expected consumption growth (ϑ_t^c) and expected inflation (ϑ_t^q), the excess nominal bond premia (over the risk-free nominal rate) are expected to change by $\rho_1\gamma\sigma_c B(\tau)$ and $\rho_2\sigma_q C(\tau)$ respectively. The differential effect of bonds of different maturity is captured by the solutions of ODEs $B(\tau)$ and $C(\tau)$ given in the appendix. Since $\lim_{\tau\to 0} B(\tau) = 0$ (same for $C(\tau)$), risk-premia vanishes for bonds that are close to maturity and increases for bonds with longer maturity. Their sign depends on the underlying parameter space - risk-aversion γ and the autoregression coefficients in the growth rate dynamics in ((3)-(4)). $B(\tau)$ and $C(\tau)$ are loadings on expected consumption growth and expected inflation in bond prices given in (13). As such, $B(\tau)$ and $C(\tau)$ should be negative, because there is a positive relationship between expected inflation and expected consumption growth and the nominal interest rate. Thus, we expect the loadings $-B(\tau)/\tau$ and $-C(\tau)/\tau$ in the yield curve equation (14) to be positive. Moreover, the sign on the risk-premia also depends on the sign of the correlations ρ_1 and ρ_2, or of the instantaneous conditional volatilities ϑ_t^c and ϑ_t^q, which determine the time-series properties of the growth rates. Since ϑ_t^c and ϑ_t^q are generated from arithmetic processes, their signs could be positive or negative. Thus, the conditional volatilities of the growth rates can produce positive or negative time-varying risk-premia. It is also noteworthy to point out that this predictability relationship is fundamentally different from the ones in ADP (2007), Ang and Piazzesi (2003), and Chernov and Bikbov (2006). In their works, the level of the nominal short-rate and the risk-premia are both driven by the same macroeconomic or latent shocks. In this case, the level of the short-rate is determined by μ_t^c and μ_t^q, whereas the risk-premia is determined by the conditional volatilities of the growth rates, ϑ_t^c and ϑ_t^q. Thus, the yield curve here has the flavor of an unspanned stochastic volatility model of the yield curve similar to Casassus, Collin-Dufresne and Goldstein (2005), but without any parameter restrictions. Cochrane and Piazzesi (2005) explore bond-return predictability with linear combinations of forward rates. They find strong evidence that information contained in the yield curve explains bond return predictability. Our approach here is similar in spirit in the sense that we are trying to extract information from the yield curve to explain premia but from an equilibrium angle. In the empirical section, we compute the right-hand side of (15) from information obtained from the yield curve and compare it against actual bond return. Notice, this formulation of risk-premia is also due to CRRA preferences. We know that CRRA preferences perform very poorly in generating high risk-premia in equities giving rise to the equity premium puzzle of Mehra and Prescott (1985). However, our focus is on bond risk-premia. The assumption of CRRA preferences greatly facilitates the closed-form solution of the yield curve and risk-premia under the non-linear dynamics of our system.

Also, notice that the real bond does not offer any risk-premia for expected inflation since it offers protection from inflation. Following Yared and Veronesi (2000), we can define inflation risk-premia to be the expected excess return on the nominal bond over the real bond, i.e. inflation-risk premia (IRP) is

$$\begin{aligned} IRP(\tau) &= \left(E_t\left[\frac{dP^n}{P^n}\right] - r_t^n\right) - \left(E_t\left[\frac{dP^r}{P^r}\right] - r_t\right) \\ &= \rho_1\gamma\sigma_c(B(\tau) - \hat{B}(\tau))\vartheta_t^c + \rho_2\sigma_q C(\tau)\vartheta_t^q \end{aligned} \tag{17}$$

To sum up, a fully structural model of the yield curve is presented here that is driven by CRRA preferences, joint conditionally Gaussian dynamics of growth rates of expected inflation and consumption growth along with time-varying conditional volatilities of expected growth-rates. Jointly, they determine the yield curve that falls into the family of quadratic term-structure models, and it produces time-varying risk-premia due to the time-varying volatilities of expected inflation and expected consumption growth. Next, we discuss a fundamental issue of macroeconomic term-structure literature that separates it from traditional factor-based term-structure models and a novel approach to estimating a joint dynamics of macroeconomic time-series and cross-section of yields.

The state-space: Before proceeding to the empirical part, let's focus on the challenges posed by the above setting. We have a non-linear nominal yield curve that is structurally determined by CRRA preferences, consumption and CPI processes ((1)-(2)), growth rate dynamics ((3)-(4)) and conditional volatilities of growth rates ((5)-(6)). The goal is to consider the full state-space - the yield curve and macroeconomic fundamentals jointly, that also preserves the parameter discipline in the yield curve induced by underlying CRRA preferences and dynamics of macroeconomic fundamentals. The full dynamics can be written jointly as

$$
\begin{aligned}
y_t^n(\tau) &= -\frac{1}{\tau}(A(\tau,\Theta) + B(\tau,\Theta)\mu_t^c + C(\tau,\Theta)\mu_t^q + D(\tau,\Theta)\vartheta_t^c + E(\tau,\Theta)\vartheta_t^q + \\
&\quad F(\tau,\Theta)\vartheta_t^{c^2} + G(\tau,\Theta)\vartheta_t^{q^2}) + \epsilon_\tau \\
\frac{dc_t}{c_t} &= \mu_t^c\, dt + \sigma_c\, dW_c \\
\frac{dq_t}{q_t} &= \mu_t^q\, dt + \sigma_q\, dW_q \\
d\mu_t^c &= (\alpha_0^c - \alpha_1^c\mu_t^c - \alpha_2^c\mu_t^q)\, dt - \vartheta_t^c\, dB_c \\
d\mu_t^q &= (\alpha_0^q - \alpha_1^q\mu_t^c - \alpha_2^q\mu_t^q)\, dt - \vartheta_t^q\, dB_q \\
d\vartheta_t^c &= (\beta_0^c - \beta_1^c\vartheta_t^c)\, dt + \epsilon_c\, dZ_c \\
d\vartheta_t^q &= (\beta_0^q - \beta_1^c\vartheta_t^q)\, dt + \epsilon_q\, dZ_q
\end{aligned}
$$

where Θ is the entire parameter space of the model. First of all, notice that there is an additional term ϵ_τ that is added to the yield equation of the state-space. It is there to break a stochastic singularity problem detailed below. Taken together - the yields and macroeconomic fundamentals, form a state-space that is at the heart of a macroeconomic term-structure *system* that should jointly address the time-series of the macroeconomic variables as well as a cross-section of yields.

Assume, there are p yields in the sample. Therefore, the full-set of observables are p yields, consumption and CPI growth for a total of $p + 2$ observables. However, there are only four latent variables. One can sacrifice the entire macroeconomic dynamics and simply use four yields to invert the state-space. But, that would defeat the purpose of this being a macroeconomic term-structure model without brining in underlying macroeconomic fundamentals. Adding the error-term ϵ_τ offers the flexibility of using more than a pre-determined set of yields and also allows us to incorporate macroeconomic dynamics as will be shown in the next section. Now using the entire set of $p+2$ observables (p yields, 2 macroeconomic fundamentals), we can *filter* the macroeconomic growth rates - μ_t^c and μ_t^q, and the corresponding conditional volatilities - ϑ_t^c and ϑ_t^q. This parameter, ϵ_τ also allows us to "mix" information from both yields and macroeconomic time-series which forms the basis of the empirical contribution of the paper and is detailed in the next section.

The growth rates of expected consumption and expected inflation enter linearly into the state-space system above. As such, we can use a standard Kalman filter to estimate the growth rates. Filtering the conditional volatilities ϑ_t^c and ϑ_t^q is a completely different problem because it enters non-linearly in the entire system - yields and growth dynamics. This paper develops a novel approach to generalize the logic of Kalman filters to produce a methodology to produce filtered *and smoothed* estimate of the conditional volatilities. The methodology is completely general and can be used in a host of models with non-linear latent shocks.

2 Empirical Methodology

2.1 Prior Choice and "mixing"

In order to see clearly the effect of ϵ_τ in this state-space system, let us write the full system in discrete time. First, define a set of variables. Let $dt = 1$ and stack the full system together. Let a τ period nominal bond yield at time t be defined by $y_t^n(\tau) = -\frac{\ln P_t^\tau}{\tau}$. Let there be p such maturities and time-series of consumption growth and CPI growth rates be denoted by $y^c = y_1^c, \cdots, y_T^c$ and $y^q = y_1^q, \cdots, y_T^q$. In our empirical methodology, we discretize the continuous-time set-up from the previous section. Following Johannes and Polson (2005) and Eraker, Johannes and Polson (2003), this form of discretization has become fairly standard in the literature. Let

$$y_{t+1} = \begin{bmatrix} y_t^n(\tau_1) \\ \vdots \\ y_t^n(\tau_n) \\ y_{t+1}^c \\ y_{t+1}^q \end{bmatrix}, \mu_t = \begin{bmatrix} \mu_t^c \\ \mu_t^q \end{bmatrix}, \, v_t = \begin{bmatrix} \vartheta_t^c \\ \vartheta_t^q \end{bmatrix} \text{ and } v_t^2 = \begin{bmatrix} (\vartheta_t^c)^2 \\ (\vartheta_t^q)^2 \end{bmatrix}.$$

Furthermore, define a set of variables

$$m_0 = \begin{bmatrix} \frac{A(\tau_1)}{-\tau_1} \\ \vdots \\ \frac{A(\tau_n)}{-\tau_n} \\ 0 \\ 0 \end{bmatrix}, \, m_1 = \begin{bmatrix} \frac{D(\tau_1)}{-\tau_1} & \frac{E(\tau_1)}{-\tau_1} \\ \vdots & \\ \frac{D(\tau_n)}{-\tau_n} & \frac{E(\tau_n)}{-\tau_n} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, m_2 = \begin{bmatrix} \frac{F(\tau_1)}{-\tau_1} & \frac{G(\tau_1)}{-\tau_1} \\ \vdots & \\ \frac{F(\tau_n)}{-\tau_n} & \frac{G(\tau_n)}{-\tau_n} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, m_3 = \begin{bmatrix} \frac{B(\tau_1)}{-\tau_1} & \frac{C(\tau_1)}{-\tau_1} \\ \vdots & \\ \frac{B(\tau_n)}{-\tau_n} & \frac{C(\tau_n)}{-\tau_n} \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$e = \begin{bmatrix} \alpha_0^c \\ \alpha_0^q \end{bmatrix}, \, f = \begin{bmatrix} \beta_0^c \\ \beta_0^q \end{bmatrix}, \, g = \begin{bmatrix} (1-\beta_1^c) & 0 \\ 0 & (1-\beta_1^q) \end{bmatrix}, \, E = \begin{bmatrix} \epsilon_c^2 & 0 \\ 0 & \epsilon_q^2 \end{bmatrix}, \, V_t = \begin{bmatrix} \vartheta_t^{c2} & 0 \\ 0 & \vartheta_t^{q2} \end{bmatrix},$$

$$W = \begin{bmatrix} \epsilon_\tau^2 I_{nXn} & 0 & 0 \\ 0 & \sigma_c^2 & \rho\sigma_c\sigma_q \\ 0 & \rho\sigma_c\sigma_q & \sigma_q^2 \end{bmatrix}, \, d = \begin{bmatrix} (1-\alpha_1^c) & -\alpha_2^c \\ -\alpha_1^q & (1-\alpha_2^q) \end{bmatrix}, \, \Sigma_t = \begin{bmatrix} 0_{nX1} & 0_{nX1} \\ \rho_1\vartheta_t^c\sigma_c & 0 \\ 0 & \rho_2\vartheta_t^q\sigma_q \end{bmatrix}.$$

and now the full state-space can be written in matrix form as

$$y_{t+1} = m_0 + m_1 v_t + m_2 v_t^2 + m_3 \mu_t + W_{t+1} \qquad W_{t+1} \sim N_{n+2}(0, W) \tag{18}$$

$$\mu_{t+1} = e + d\mu_t + U_{t+1} \qquad U_{t+1} \sim N_2(0, V_t) \tag{19}$$

$$v_{t+1} = f + g v_t + E_{t+1} \qquad E_{t+1} \sim N_2(0, E) \tag{20}$$

and Σ_t is the covariance between the innovations in (18) and (19). (18) is the "observation" equation in the sense that we observe the yields and the macroeconomic series of consumption and CPI. Equations (19)-(20) are the "state" equations in the sense that they describe the evolution of the underlying latent states. All the parameter restrictions are embedded through the no-arbitrage ODE's stacked in the coefficient on the yield curves in $\{m_0, m_1, m_2, m_3\}$. Furthermore, notice the non-linear nature of the state-space - v_t enters non-linearly in the observation equation (18) as well as in the state equation (19).

To emphasize the role of priors, notice that the theory does not guide us to determine how to tie down ϵ_τ. Thus far, the role of ϵ_τ is to break stochastic singularity but it can also be used to estimate parameters and latent-variables by mixing information from both the yield curve and macroeconomic fundamentals. The observation equation (18) reveals that if one were to set the prior on $\epsilon_\tau$in such a way that the posterior distribution of $\epsilon_\tau$is centered on a very small number, then the resulting state-variables and parameters will be very strongly influenced by the yield curve and less by the macroeconomic fundamentals. If, however, one were to set the prior on $\epsilon_\tau$such that the posterior is centered on a relatively larger number, then the same resulting parameters and state-variables will be determined primarily by the macroeconomic fundamentals and less by the yield curve. A prior setting in between on $\epsilon_\tau$will extract information - i.e. estimate state-variables and parameters, from both yield curve and the macroeconomic fundamentals.

Notice that this analysis embeds three different types of term-structure models. The **tight** prior which draws information from only the yield curve is isomorphic to a four-factor Gaussian term-structure model that gives rise to a quadratic yield curve. As will be shown, it does what factor models are designed to do - it fits the yield data tightly and draws state-variables and parameter estimates from the yield curve but forsakes the macroeconomic series. As such, it is no-different from a traditional factor model where the factors are statistically chosen without much economic content. The **loose** prior primarily draws parameter estimates from the macroeconomic data. It results in a poor fit of the yield curve as the **loose** prior fails to extract substantial information from asset prices. The **joint** prior draws information from both the yield curve and the macroeconomic time-series.

Different from a statistical model whose goal is to tightly fit the yield curve - like the **tight** prior, this prior setting draws information from both the yield curve and the macroeconomic series thereby giving it a strong macro term-structure flavor. The resulting fits with the yield curve are not as good as the **tight** setting, but substantially better than the **loose** setting. Moreover, the state-variables extracted from this setting looks much closer to the economic growth rates and volatilities than the **loose** setting. This way of looking at the problem is novel in the literature since the theory simply does not guide us on the size of this pricing error-term.

Another interesting observation from the **joint** setting is that the risk-premia generated from this setting does a much better job in explaining bond returns than the **tight** or the **loose** setting. Taking the distribution of parameters and the latent state variables, we construct the time-series of nominal bond risk-premia according to (15) and match it against bond return data constructed from the same yields we are matching empirically. We find that the parameters and state-variables extracted from the **joint** prior that pulls in information from both yield curve and macroeconomic fundamentals does a superior job in bond return predictability over the **tight** setting that fits the yield curve very closely. Details are provided in Section 3. This displays the strength of macroeconomic models over statistical ones. Whereas the latter does a superior job in matching a *sample* of yields, it fails to live upto other model implications, like risk-premia that can address predictability. The strength of this paper is that all of the above is shown in a very restrictive equilibrium framework where preferences and dynamics fully tie-down the yield curve, and does not have the additional degrees of freedom from having an exogenous parameterization of the risk-premia as in ADP (2007) or CSV (2008).

2.2 Bayesian MCMC

We use a Bayesian Markov Chain Monte Carlo (MCMC) methodology - the Metropolis-Hastings sampler, to make inference about the parameter space and the latent states. Bayesian methodologies to estimate dynamic models in finance have been made popular through Johannes and Polson (2005), Eraker (2004) and others. They employ similar Metropolis algorithms, but the non-linear filtering methodology is novel to this paper that generalizes the logic of the Bayesian version of Kalman filter called Forward Filtering Backward Sampling and is explained in the Appendix. Thus far, filtering stochastic volatility in the Bayesian literature goes back to Jacquier, Polson and Rossi (1994) and Kim, Shepherd and Chib (1998). The main difference between this paper and Jacquier, et. al. (1994) is that the conditional volatilities here are both filtered and smoothed which avoids the high autocorrelation in the states of Jacquier, et. al.(1994). The filtering technique of Kim, et. al. (1998) is of a very particular parameteric model which does not correspond to the model presented here. CSV (2008) also employs a non-linear filtering mechanism. The difference between their methodology and ours is that ours is based on the full probability distribution of the space we are trying to filter from conditional on the data and the rest of the parameters and growth rates, whereas their methodology relies on the local approximation of the first two moments. Secondly, our methodology fits into the MCMC paradigm providing an unifying framework to filter and estimate parameters jointly.

Very briefly, the Bayesian methodology gives the joint distribution of

$$p(\mu, v, \Theta|y)$$

where the full data sample is denoted by y that includes a cross-section of the yield curve along with the macroeconomic dynamics, $\mu = \{\mu_1, \cdots, \mu_T\}$ the expected growth rates, $v = \{v_1, \cdots, v_T\}$ the conditional volatilities of the growth rates and Θ is the full set of parameters of macro quantities and preferences. The Metropolis-Hastings sampler is used to draw the parameters and the state-variables jointly by drawing them one conditional on the other - first the state-variables and then the parameters.

$$
\begin{aligned}
&p(\mu|v,\Theta,y)\\
&p(v|\mu,\Theta,y)\\
&p(\Theta|\mu,v,y)
\end{aligned}
$$

They can be further broken down to draw $p(\{\vartheta^q\}|\{\vartheta^c\},\cdots)$ or $p(\Theta_i|\Theta_{-i},\cdots)$, where Θ_{-i} is the set of Θ modulo the i-th element. Notice that in the presence of the yield curve in the observation equation, we do not have closed form posterior distribution for any parameter. For example, for the i-th parameter, say the regression coefficients (e,d) in (19), the posterior distribution looks like

$$p(e,d|\Theta_{-(e,d)},\mu,v,y)=p(e,d)p(\mu|(e,d),v,\Theta_{-(e,d)})p(y|(e,d),\Theta_{-(e,d)},\mu,v,y)$$

The first term in the posterior distribution is the prior on the regression coefficient whereas the second term is the likelihood conditional on the expected growth rates, μ. The third term is the likelihood conditional on the yield curve. Notice that the parameters (e,d) enter into the yield equation very non-linearly through the no-arbitrage restriction in the model imposed in $\{m_0,\cdots,m_3\}$ in (18). The effect of the prior distribution on the error term of the yields - ϵ_τ is now clear. Under the **tight** setting, the posterior distribution of (e,d), or on any parameter, will be strongly influenced by the likelihood of yield curve. Under the **loose** setting, the posterior distribution of (e,d) will be determined by the time-series of the macroeconomic fundamentals.

There is another important issue in the inference of these parameters. Without the bond equation - i.e., if the state-space was estimated without the yield curve, then closed-form solutions are available. In the presence of the yield curve, no known posterior distribution (or proposal) of the parameters is available. As such, using random-walk Metropolis, as is frequently done in the finance literature, may not always "work" for all parameters. See Chib and Ergashev (2008). For some parameters, like the regression coefficients, the random-walk Metropolis works fine, but not for others like the volatility of volatilities ϵ_c and ϵ_q, the correlations ρ_1 and ρ_2, or the preference parameters. In these instances, we implement a Gibbs sampler on a grid. For the volatility of volatilities, the grid is chosen based on the posterior distributions of these parameters from the macroeconomic data, whereas for the correlations the grid is $(-0.99, 0.99)$. For the preference parameters (risk-aversion and time-preference), we impose the left boundary at 0 because they cannot be negative. The details of the estimation methodology is provided in Appendix B.

The next step is to draw the growth rates $\mu|v, \Theta, y$ which is a linear filtering problem since the growth rates μ enter linearly in the yield curve and macroeconomic growth rates (18) and their transition densities are also linear (19). The Bayesian version of Kalman filter that is used to draw the growth rates is called Forward Filtering Backward Sampling (FFBS) that is due to Carter and Kohn (1994) and Fruhwirth-Schnatter (1994), and is also used by ADP (2007). A small distinction here is the correlation in the error terms between the observation (18) and state equation (19). We resolve it by orthogonalizing (19) by using the fact that $U_{t+1}|W_{t+1}, v$ is still normally distributed, and then the standard FFBS filter with independent error terms follow. Finally, we filter $v|\mu, \Theta, y$ which is a non-linear filtering exercise since the conditional volatilities enter non-linearly in both the yield curve and the macroeconomic dynamics. In order to filter these conditional volatilities, we generalize the logic of FFBS used for the growth rates to a non-linear setting. Details are available in the Appendix.

2.3 Data and Prior Specification on ϵ_τ

The data for the empirical part comes from the post-war US aggregate macroeconomic series from 1959-2005 sampled monthly. Aggregate per-capita real consumption data comes from Bureau of Labor Statistics (non-durables and services) divided through by a population series to make it per-capita consumption. US CPI data comes from International Financial Statistics. Yield data comes from the Fama-Bliss file obtained from CRSP on 1-5 year bonds.

Notice that the observation equation of this state-space dynamics contains both yields and macroeconomic time-series. As a first-pass, we first estimate the macroeconomic state-space alone without the yields. This allows us to form valuable prior opinion on the entire parameter space except for the preference parameters. In this case, all the posterior distributions of parameters are known in closed form, and one can draw them with computational ease. Then, these posteriors can be used as priors when we bring in the full state-space for which no known posteriors (or proposals!) are known.

To gauge the prior on the observation error on the yields - ϵ_τ, the data on yields provide us some guidance on what would be considered a "tight" or "loose" fit. The standard deviation on all monthly yields is roughly 22 basis points. Given this fact, we motivate our prior on ϵ_τ the following way. The prior on ϵ_τ is chosen to be inverted chi-square with hyper-parameters such that the posterior distribution of ϵ_τ is centered at 0.0003^2 in the **tight** setting, 0.0020^2 for **loose** setting and 0.0010^2 for the **joint** setting. The hyper parameters are also picked in a way such that the distribution of ϵ_τ in each case is tight around their respective means.

The **tight** prior on ϵ_τ seeks to fit the yield curve very tightly a la a statistical 4-factor Gaussian model within a 3 basis point standard deviation. Clearly, this prior specification will draw parameters and latent variables that will correspond to a very tight fit with the yield curve mimicking a statistical factor model. On the other end, the **loose** prior setting fits the yield curve very loosely and draws parameters and latent variables mostly out of the macroeconomic fundamentals. Thus, this prior setting essentially forsakes information from the yields and seeks to fit the macroeconomic fundamentals. The setting of most interest - the "joint" prior, seeks to mix information from the yield curve and the macroeconomic fundamentals jointly. Since it is not fitting the yield curve too tightly this prior setting can elicit information from both yields and the macroeconomic fundamentals.

In Bayesian parlance, the role of ϵ_τ is to provide shrinkage. Under the **tight** prior setting, the state-space *shrinks* towards the yield curve and provides inference on parameters and state variables based on the yield curve. Under the **loose** prior setting, the system shrinks towards the macroeconomic fundamentals, and elicits parameters and state-variables from the macroeconomic fundamentals. The joint setting shrinks towards both and pulls in information from both the yield curve and the macroeconomic fundamentals. This framework is currently missing in the macro-term structure literature. At the heart of any macroeconomic yield curve, there is both macro-dynamics and an yield curve that is derived from it based on no-arbitrage. There is a clear tension between parameters/state-variables that drive macroeconomic time-series *and* the corresponding yield curve. As such, any empirical study must address this tension formally and this paper seeks to understand what happens if we start to *shrink* from a statistical model of the yield curve (the **tight** prior restriction) to a setting that brings in macroeconomic fundamentals (the **joint** prior restriction).

3 Empirical Findings

In order to gauge the parameters of the macroeconomic time-series, we estimate the state-space without bonds and without any prior information. This gives us indication on where the parameters of the underlying macro dynamics lie. Then we bring in the yield curve, and execute the three different prior settings. For each prior specification of the Metropolis-Hastings algorithm, we run the markov chain for 25000 iterations and discard the first 15000 for burn-in, and form our posterior distribution about parameters and latent variables based on the remaining 10000.

The *No Bonds* column of Table I gives us indication of what the underlying parameter space looks like without any prior information and without the yield curve. This gives us good a good prior distribution of parameters for full analysis with the yield curve. Closed-form posterior distributions are available for the *No Bonds* case facilitating the task of obtaining posterior draws of the parameter space. The off-diagonal terms of the growth rate VARs reveal that there is a strong feedback effect. High expected inflation lowers future expected consumption growth rate, whereas the effect of high growth rate on future expected inflation is inconclusive as the estimate of α_q^1 is large but has a wide posterior interval. The conditional volatilities are fairly autoregressive with the autocorrelation parameter being 0.8551(0.8567) for conditional volatility to expected consumption growth (expected inflation). The long-run average of the conditional volatility in expected consumption growth rate is $\frac{\beta_c^0}{1-(1-\beta_c^1)} = \frac{0.0004}{1-0.8551} = 0.0027$ whereas the average conditional volatility in expected inflation is $\frac{\beta_q^0}{1-(1-\beta_q^1)} = \frac{0.0002}{1-0.8567} = 0.0014$. In the sample, the volatility of monthly consumption growth is 0.0036 and the volatility of monthly CPI growth is 0.0031. Thus, the volatility of the *expected growth rates* of both quantities - consumption and CPI, is smaller showing that the growth rates are substantially smoother - especially for inflation, than the observed series. Furthermore, the volatility of the conditional volatilities are $\epsilon_c = 0.0014$ and $\epsilon_q = 0.0009$ showing that the conditional volatilities of the growth rates are further incredibly smooth. This is expected from aggregate macroeconomic series, even if they are sampled at a monthly frequency. Now let us augment the state-space and add the structurally determined yield curve and see how these parameter estimates and state variables change.

First, we look at the **loose** setting. Naturally, under this prior the parameters and the state-variables that are estimated will not be strongly influenced by the yield curve but by the macroeconomic variables. In Bayesian parlance, the parameters *shrink* more towards the macroeconomic growth rates than towards the yield curve. Naturally, the parameter estimates do not differ very much from the *No Bonds* setting. The only major distinction is that the volatility of consumption and price growth are somewhat smaller compared to the *No Bonds* case suggesting that the state-variables and parameters fit the macroeconomic series very tightly. Figure 1 and Figure 3 present the growth rate state-variables $\{\mu^c\}$ and $\{\mu^q\}$ that are obtained from the loose prior setting using FFBS. The median fits are very tight with the underlying macroeconomic variables, and the 2.5-97.5 quantiles do not depart much further from the data. But, this is to be expected. The **loose** prior setting was designed not to "shrink" towards the yield curve, but towards the macroeconomic series. Therefore, the underlying parameters imply states that match the macroeconomic series very closely. The conditional volatilities of these growth rates are shown in Figures 5 and 6. They are essentially constant! The conditional volatilities of the growth rates that are implied by the macroeconomic variables are minute in size and essentially non-stochastic. This is a good test that our non-linear filtering methodology to extract the conditional volatilities is "working" because we expect these macroeconomic volatilities to be small and stable. As expected, the fits with the yield curve are poor. Figures 9-11 show that the state-variables and parameters obtained from the **loose** settings provide very poor fit with the time-series of yields. The dotted lines in Figures 9-11 show that the time-series of yield curve implied by the state-variables and parameters of the **loose** setting completely fail to match the time variation in the yield curve. Whereas the median posterior estimate goes through the center of the data, the setting here simply cannot generate enough volatility to capture the time-series variability of the yields. Table III shows the posterior distribution of pricing errors for the **loose** setting. For monthly yields, the **loose** setting is off on average by 16 basis points for the 1-year yield to 14-basis points for the 5-year yield, which is quite a large error on monthly yields. To sum up, the state-variables and parameters that capture the time-series of the macroeconomic variables simply cannot generate the variability needed to capture the time-series variation in observed yields. The conditional volatilities of the growth rates - which were responsible for the non-linearity in the yield curve and generating time-varying risk-premia, simply isn't volatile enough to create any meaningful time-series variation in yields.

The **tight** setting shifts our attention to the bonds. It fits the yield curve within a spread of 3 basis points, which given the standard deviation of the yields, is an attempt to fit the yield curve very tightly. The macroeconomic series are still part of the state-space, but since this prior setting puts so much emphasis on the yield curve, the parameters and the state variables are influenced strongly by the latter. Again, figures 9-11 show how the **tight** setting does in fitting the yield curve over the **loose** setting. The fit is almost perfect! The underlying parameters and the state-variables imply an yield curve that lines up almost perfectly with the data. The posterior median fit is almost flawless as the underlying state-variables generate all the variability that is needed in order to fit the yield curve. The median pricing error for the **tight** setting is between 3 basis points for the 2-year yield to 1.5 basis points for the 1-year yield suggesting a near perfect match. But, at what cost?

This astronomical improvement in fitting the yield curve comes at a substantial price with regards to fitting the underlying macroeconomic series. The bottom panels of Figures 2 and 4 show the posterior distribution of the growth rates that were supposed to have been expected consumption growth and expected inflation. Both fail dismally in bearing any resemblance to the underlying macro series. The bottom panel of Figure 2 produces what should have been filtered and smoothed estimates of expected consumption growth. The filtered estimates look anything but that. In fact, the time-series of this “factor” - high in the early 80s, stable in the 90s, and low and negative between 2000-2005, suggests that the yield curve is artificially creating a factor that has time-series properties of US inflation in order to fit the yield curve. Figure 4 is even more radically different. What is supposed to be filtered and smoothed estimates of expected inflation doesn't even come close to it. In fact, the time-series of the second growth rate “factor” is simply something that is negatively correlated with price growth and is the opposite of expected inflation. Similarly, Figures 5 and 6 show the posterior median of the conditional volatilities of expected consumption growth and expected inflation. The tight fit with the yield curve implies that the corresponding conditional volatilities are large with huge time-series variation that changes sign, and most importantly, bear no resemblance with the almost non-stochastic conditional volatilities that the macro series imply in the **loose** setting. This is yet another test of the non-linear filtering technique. We know from the **loose** setting that we need state variables that are volatile in order to match the yields. What helps us to fit the yields are these conditional volatilities that enter non-linearly and are highly volatile.

What is happening is endemic to all reduced form term-structure models. One can invert the yield curve in these pure factor models and can produce a tight fit and generate a time-series of factors. However, these factors are not easily interpretable and not tied-down to any economic fundamentals. For example, the yield curve in (14) can be interpreted as a 4-factor non-linear term structure model where all the factors are Gaussian. What is produced through the **tight** prior setting is a strong fit with the yield curve a la a reduced form setting which forsakes all relationship to a macroeconomic term-structure model that also needs to address the underlying economic setting of macro dynamics and preferences of agents exposed to these shocks. The **tight** setting pays no attention to the latter and simply generates a time-series of "factors" to do what factor models do best, i.e. generate fits. But, the corresponding factors - in this case expected inflation and expected consumption growth along with the stochastic conditional volatilities of these variables, do not look anything like the latent factor dynamics extracted from the **loose** setting. The only conclusion we can draw from the **tight** setting is that it produces a strong fit but to be able to do so, the corresponding growth rates have no bearing with the underlying economic setting.

The setting in between the **loose** and the **tight** is the **joint** setting which attempts to shrink towards both the macroeconomic fundamentals and the yield curve, and in doing so, tries to extract information from both. The median pricing error in this case is 10 basis points, which is not as strong as 3 basis points for the **tight** case, but not as loose as 20 basis points for the **loose** prior case. Thus, the fits with the yield curve that this prior specification generates is in between the **loose** and the **tight** case, but definitely closer to the **tight** . Figures 9-11 show that even if the fits are not as close as the **tight** prior setting, they are substantially better than the loose setting. The fits in the **joint** setting capture all the time-series variation in the underlying yield curve, but do not fit them as closely as in the **tight** setting. The median pricing errors in the **joint** setting range from 7.6 basis points for the 1-year yield to 3.5 basis points for the 5-year yields. Again, the pricing errors are not as small as the **tight** setting, but clearly better than a factor of 2 over the pricing error of the **loose** setting.

At the same time, this setting provides substantial better fit with the macroeconomic data. Figures 2 and 4 show the same time-series of expected consumption growth and expected inflation from the **joint** setting in the top panel. The time-series of the growth rates reveal that the **joint** setting does a much better job in matching the time-series properties of the underlying macro series than the **tight** setting. It produces expected consumption growth and expected inflation that are much closer to their realized counterpart than the **tight** setting. Naturally, the fits with the macroeconomic series is not as strong as the **loose** setting, but they are substantially closer than the **tight** setting. The time-series of the conditional volatilities also show shrinkage towards the macroeconomic series. Figures 5 and 6 show that the conditional volatilities of the growth rates are much closer to the stable conditional volatilities of the **loose** setting than the wildly fluctuating ones of the **tight** setting.

An important caveat is in order. Even if the **joint** setting provides a superior fit with both the yield curve and macroeconomic series, it seems that there is a clear structural break in the early 1980s. To get closer to the high yields of the early 1980s, expected inflation and expected consumption growth deviate substantially from the underlying macro series. Similarly, from 2000-2005, the **joint** setting essentially produces expected deflation in order to track the low yields in that period. In terms of the time-series of the stochastic volatilities, the conditional volatility of expected inflation increased substantially in the early 1980s (Figure 6) and dropped soon after suggesting that the structural break in expected inflation resulted from the sudden increase in conditional volatility of expected inflation which is needed to track the high yields of the early 80s.

Overall, however, the **joint** prior setting - which draws information from both the macroeconomic data and the yield curve performs well in matching both yields and the underlying macroeconomic fundamentals. The **joint** setting illustrates the challenge of any macroeconomic yield curve *system* where the goal is to draw joint inferences on macro fundamentals and the yield curve that are connected by preferences and/or no-arbitrage restrictions. The setting here addresses this point with the choice of priors to elicit information from one or the other, or both. As we show here, the prior on the **joint** setting that draws information from the two different series can accomplish both - it produces substantially better fit with the yield curve than the **loose** setting, and substantially better fit with the macroeconomic growth rates than the **tight** setting.

3.1 Risk Premia:

Now we show where mixing information from both the yield curve and macroeconomic fundamentals is most valuable. Taking the posterior parameter distributions and state variables, we construct the risk and term premia offered on bonds. We show that the **joint** setting which elicits information from both the yield curve and the macroeconomic fundamentals is able to produce premia that is better able to address predictability than the **tight** or **loose** setting. From the data on monthly yields, we construct the return data by taking a τ-period bond and sell it one year later as a $\tau - 1$-period bond. We construct excess return by subtracting the yield on the 1-year bond whose rate of return is immediate. At the same time, we construct the time-series of the posterior distribution of risk-premia for each bond with maturity τ following the expected excess bond return expression

$$(\rho_1 \gamma \sigma_c B(\tau) \vartheta_t^c + \rho_2 \sigma_q C(\tau) \vartheta_t^q) \cdot 12$$

We multiply it by 12 in order to create annual estimates since our return data is annual.

The time-series of the median risk-premia for each bond from both the **tight** and **joint** setting is reported in Figures 12-13. First, the risk-premia from the **tight** setting is simply too high and too erratic. It shows risk-premia steadily increasing from the 1960s onward until risk-premia reaches almost as high as 25% for the 5-year bond and 15% for the 4-year bond in the late 1970s. Then, dramatically, the risk-premia changes sign reaching well below -25% on the 5-year bond and slowly coming back up in the later part of the sample. The **tight** setting fits the yield curve very tightly, but generates risk-premia that are simply too high. In fact, the premia in this case is strongly *negatively* correlated with observed return. Table IV shows that for the **tight** setting, the correlation between the risk-premia shown in Figure 12 and the observed excess return on the bond is as high as -0.56 for the return on a 5-year bond to -.60 for the return on a 2-year bond. The **tight** setting was very good in terms of producing fits, but it failed completely in living up to other model implications like predictability. However, the time-series of risk-premia shows a very interesting pattern. It is exclusively positive and rising in the pre-Volcker regime. The early 1980's with new Fed policy completely reverses the sign on the premia and it has slowly increased since then.

The results from the **joint** setting are more promising. Constructing the same time-series of premia from above using the posterior distribution of parameters and state-variables from the **joint** setting, we produce the risk and term premia in Figure 13. We notice that the risk-premia is much smaller than the **tight** setting, but it correlates positively with the observed return. The third column of Table IV shows that the time-series of risk and term premia implied by the parameters and state-variables of the **joint** setting produces correlation of 0.19 with the return on a 2-year bond to 0.22 with the return on 4-year bond. Moreover, the risk-premia from the **joint** setting can also generate economically significant R^2's. Table V shows that when the bond excess return is regressed against the risk-premia generated from the **joint** setting, the median R^2's that are generated range from 3.5-5% with their posterior distribution rising as high as 8-10% for bond returns of maturity 2-5 years. We compare our R^2's against ADP (2007) and the full posterior distribution is well within reach of the R^2's that are generated in their work. This is significant due to the strong parameter restrictions imposed in our setting that ties down the risk-premia coefficients to preferences and macroeconomic dynamics.

To understand the different predictability results across the different prior settings, let's regress the conditional volatilities of expected inflation ϑ_t^q and expected consumption growth ϑ_t^c that we filtered from the yield curve to bond excess returns. The coefficients are shown in Table VI. Changes in conditional volatility of expected inflation is strongly negatively correlated with excess bond return in both the **joint** and **tight** setting and the effect is larger for higher maturities. However, the predictability coefficients are not free parameters for us, since in (15) we have pinned down these coefficients through preferences, underlying dynamics and no-arbitrage. Only the **joint** setting can match the sign and pattern of the coefficient $(\rho_2 C(\tau)\sigma_q)$ on ϑ_t^q that is presented alongside the regression coefficients.[3] They are negative and decreasing for longer maturities implying that a positive shock to conditional volatility of expected inflation predicts bond prices will fall, and longer maturity bonds are expected to have a lower expected excess return as a result of increased volatility than shorter maturity bonds. The same coefficient that we compute in the **tight** setting gives us the wrong sign and the reverse inference - higher conditional volatility to expected inflation is supposed to increase expected return and longer maturities have higher expected excess return than shorter maturities. The effect of conditional volatility of consumption growth on expected bond excess returns is not conclusive since the coefficient is insignificant in the **joint** setting and is negatively correlated with bond returns in the **tight** setting, but the effect is not as strong as conditional volatility of expected inflation.[4]

The reason the **joint** setting gets the correct sign on the predictability relationship is because it gets the right sign on the loading $C(\tau)$ - which was the coefficient on expected

[3]These values are simply taken from Table II.

[4]In these regressions, the R^2's are substantially higher than the R^2's in Table 5, but we do not report them since these regressions do not enforce the parameter restrictions which is the purpose of our study.

inflation in the bond pricing relationship (13). As such, we should expect $C(\tau)$ to be negative, or $-C(\tau)/\tau$, i.e. the coefficient on expected inflation in the yield curve, to be positive. The bottom panel of Figure 16 confirms that hypothesis. It also confirms that the **tight** setting gets the wrong sign on $C(\tau)$. The reason the two settings get two different signs on the loading has to do with how much each shrinks towards the macroeconomic setting. Since the **joint** setting shrinks substantially to the macroeconomic variables, it is able to filter expected inflation from CPI data. On the other hand, since the **tight** setting disregards the macroeconomic series completely, the expected inflation series it produces is completely different (and, in fact, orthogonal to expected inflation). That is essentially shown in Figure 4. Since the **joint** setting is better able to identify expected inflation, the corresponding loading it generates on expected inflation ($-C(\tau)/\tau$) from the yield curve relationship (14) is positive and increases with maturity as is shown in the bottom panel of Figure 16. Since the **tight** setting cannot identify expected inflation from CPI data (it disregards information from the macro series), it doesn't get the right sign on the loading of this factor and eventually on premia. Precisely which underlying variable is causing this phenomenon is difficult to address, but as the ODE's satisfying the loadings show in the Appendix, it is the autocorrelation parameters of the growth rates and risk-aversion which are involved in producing this differential effect.

The fact that the **loose** setting is not able to produce any conclusive relationship between returns and premia is because the macroeconomic series is very poor in identifying the correlation parameters ρ_1 and ρ_2 which govern the correlation between expected consumption growth (expected inflation) and observed consumption (CPI) growth. The **loose** setting infers that these parameters are indistinguishable from zero, and furthermore the latent conditional volatilities are virtually constant. Hence the correlation of premia extracted from the **loose** setting with actual return is essentially zero.

Another caveat is in order. Whereas the **joint** setting can produce premia that correlates with actual bond returns, it generates premia that is small in magnitude. At most, the median premia is .5%, and it can generate a negative premia as low as -5%. However, small and negative premia is not unexpected. In fact, Veronesi and Yared (2000) also show in their equilibrium setting small and negative term premia before their "inflation convexity bias" correction due to learning. We can generate substantial premia from the **tight** setting, but it gets the opposite sign in the predictability relationship and is not in spirit of a macroeconomic term-structure model because it disregards the underlying macroeconomic information.

3.2 The effect of non-linearity on yields

The effect of non-linearity is severe in the conditional volatility of expected inflation, particularly for the long maturity bonds. The bottom panel of Figure 18 shows the loading on the $\vartheta_t^{q^2}$ term. If the size of the conditional volatility of expected inflation is big (sign doesn't matter here), then long maturity yields suffer more than short maturity yields. However, the effect of the non-linearity of the conditional volatility to expected consumption growth is much smaller in magnitude and its effect on the yield curve is not monotonic.

3.3 Parameter Estimation:

Ultimately, the three settings are distinguished by the parameters of the underlying dynamics and preferences (like risk-aversion). The **tight** setting, which ignores the macroeconomic variables, estimates the volatility of consumption growth, σ_c, to be 0.0159 in *monthly consumption.* That amounts to $0.0159 \cdot \sqrt{12} = 5.5\%$ volatility of yearly consumption growth which clearly indicates that the **tight** setting absolutely disregards the underlying macro series. The volatility of price growth does even poorly with a monthly estimate of σ_q to be 0.08. Naturally, the premia generated from the **tight** setting is much higher than the other two settings.

The correlation parameter ρ_1 (ρ_2) that measures correlation between expected consumption growth (expectation inflation) and observed consumption (CPI) growth are very different under the yield curve dynamics than the macro dynamics. For example, the macroeconomic series offer no conclusive inference about these variables since they are both indistinguishable from zero. However, the yield curve is very informative about these variables. Under the **tight** (or **joint**) setting ρ_1 (ρ_2) is strictly negative (positive) implying a negative (positive) relationship between consumption growth (price growth) and expected consumption growth (expected inflation) *as inferred by the yield curve.* The reason the yield curve is informative about these parameters is because the **tight** (or **joint**) setting is looking to extract premia from the yield curve in order to generate fits with the yield curve. One way to infer premia is to strongly tie down these correlations which the prior setting is helping to do. One can think of the three different settings as different implications of the failure of Expectations Hypothesis. The **loose** setting that extracts no statistically significant premia displays the features of the Expectations Hypothesis, but it cannot match the yields. The **tight** setting fits the yields but produces an extreme rejection of the Expectations Hypothesis. The **joint** setting fits the yields reasonably well, and it also extracts reasonable premia showing the failure of Expectations Hypothesis.

The effect of the prior is also visible through the estimation of the risk-aversion parameter γ. Figure 15 plots the posterior distribution of γ under all three prior setting. The **tight** setting implies that posterior distribution of risk-aversion is between 4.5-5, which is again one of the reasons why premia is much higher for this prior setting. At the same time, the **loose** setting is reporting risk-aversion much lower between 0.25-0.75 precisely because it is not getting enough information from bond prices due to the loose prior that we imposed. The **joint** setting which draws moderate amount of information from bonds infers γ between 3.5-4.

As discussed above, one can also test the Fisher hypothesis of independence of real interest rates and expected inflation. In this setting $\alpha_2^c \neq 0$ implies that the long run mean of the real rate is determined by expected inflation. The top panel of Figure 14 shows that the posterior distribution of α_2^c is negative, suggesting that high expected inflation increases expected consumption growth which positively affects the real rate. This puts it at odds with the Mundell-Tobin effect. The impact of adding information from the yields on the same parameter is obvious in the bottom panel. Clearly the presence of the yields tighten the parameter space. Whereas the posterior distribution of α_2^c without the yield is fairly wide, the introduction of the yield curve is very informative on where the parameter should lie and pins it within a small sub-interval of the posterior distribution of the *No Bonds* case. Whereas there is a small probability in the *No Bonds* case for Mundell-Tobin effect to hold (positive α_2^c), the information from yields make it conclusive that this parameter is negative.

3.4 Inflation Risk-Premia:

Following Veronesi and Yared (2000), we define inflation risk and term premia to be the difference between expected excess return of a nominal bond and expected excess return of a real bond. It is given by the expression

$$IRP(\tau) = \rho_1 \gamma \sigma_c (B(\tau) - \hat{B}(\tau)) \vartheta_t^c + \rho_2 \sigma_q C(\tau) \vartheta_t^q$$

Taking the full posterior distribution of parameters and latent conditional volatilities from the **joint** prior setting, we construct the posterior distribution of inflation risk premia across the different maturities. We report the median time-series estimates in Figure 19. Just like risk-premia itself, we find that the magnitude of inflation risk premia across the different maturities is small and mostly negative. The exact sign and magnitude of inflation risk premia is widely debated in the literature. Negative and tiny inflation risk-premia has been uncovered in Ravenna and Seppala (2007) and Veronesi and Yared (2000), whereas Buraschi and Jiltsov (2005) uncover high inflation risk-premia as an important explanatory variable for deviations from the expectations hypothesis. In our case, the **joint** setting implies that the inflation risk-premia has been fairly constant in the time-series. It is mostly negative and the early 1980s saw a big negative jump in the inflation risk-premia. The evidence suggests that nominal bonds offer no reward for agents exposed to inflation risk. It is most likely the case that nominal bonds are beneficial for the agent to offer diversification benefit in his/her extended portfolio.

4 Conclusion

We propose a a macroeconomic no-arbitrage yield curve *system* that seeks to address both no-arbitrage yields with underlying macroeconomic fundamentals connected by preferences that puts further restriction between macro dynamics and risk-premia. We estimate a joint state-space of the quadratic yield curve with underlying macroeconomic fundamentals using a Bayesian MCMC algorithm that estimates parameters and state-variables by using prior information to elicit information from either the macroeconomic growth rates or the yield curve, or both. We conclude that effectively using priors to elicit information out of both yields and macroeconomic fundamentals have the advantage that it can explain bond return predictability much better than if we only looked at yields or macroeconomic fundamentals separately. Our analysis is in partial equilibrium, thereby extending the macroeconomic term-structure literature by imposing restriction on the risk-premia that provides the connection between macro dynamics and no-arbitrage bond prices under failure of Expectations Hypothesis.

References

[1] Ang, A., S. Dong and M. Piazzesi, 2007, No-Arbitrage Taylor Rules,*Working Paper*.

[2] Ahn, D.-H., R.F.Dittmar and A.R.Gallant, 2002, Quadratic Term-Structure Models: Theory and Evidence, *Review of Financial Studies*, 15, 243-288.

[3] Ang, A. and M. Piazzesi, 2003, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monetary Economics*, 50:4, 745-787.

[4] Beaglehole, D. and M. Tenney, 1991, General Solutions of Some Interest Rate Contingent Claim Pricing Equations, *Journal of Fixed Income*, 6, 43-53.

[5] Bernanke, B.S. and I. Mohov, 1998a, The liquidity effect and long-run neutrality, *Carnegie-Rochester Conference Series on Public Policy*, Vol 49, B.T. McCullum and C.I. Plosser(eds.) Amsterdam: North Holland, 149-194.

[6] - and -, 1998b, Measuring monetary policy, *Quarterly Journal of Economics*, 113, 869-902.

[7] Bikbov, R. and M. Chernov, 2005, No-Arbitrage Macroeconomic Determinants of the Yield Curve, *Working paper*.

[8] Buraschi, Andrea and Alexei Jiltsov, 2007, Habit Formation and Macroeconomic Models of the Term Structure of Interest Rates, *Journal of Finance*, 62, 3009-3063.

[9] 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.

[10] Campbell, John Y., Adi Sunderam and Luis M.Viceira, 2008, Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds, *Working Paper*.

[11] Carter, C.K. and R. Kohn, 1994, On Gibbs Sampling for State Space Models, *Biometrika*, 81, 541-553.

[12] Casassus, J. P. Collin-Dufresne, R. Goldstein, 2005, Unspanned stochastic volatility and fixed income derivatives pricing, *Journal of Banking and Finance*, 29, 2723-2749.

[13] Clarida, R., J. Gali and M. Gertler, 1998, Monetary Policy Rules in Practice: Some International Evidence, *European Economic Review*, 42:6, 1033-1067.

[14] Clarida, R., J. Gali and M. Gertler, 2000, Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory, *Quarterly Journal of Economics*, 115, 147-180.

[15] Cochrane, J. and M. Piazzesi, 2005, Bond Risk Premia, *American Economic Review*, 42:6, 1033-1067.

[16] Cogley, T. and T. Sargent, 2001, Evolving Post-World War II U.S. Inflation Dynamics, *NBER Macroeconomics Annual*, 16, 331-373.

[17] Constantinides, G., 1992, A Theory of the Nominal Structure of Interest Rates, *Review of Financial Studies*, 5, 532-552.

[18] Dai, Q. and K.J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance*, 55, 1943-1978.

[19] Dai, Q. and K.J. Singleton, 2002, Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure, *Journal of Financial Economics*, 63, 415-441.

[20] Diebold, F.X., G.D. Rudebusch and B. Aruoba, 2006, The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach, *Journal of Econometrics*, 131, 309-338.

[21] Duffee, G.R., 2000, Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance*, 57, 405-443.

[22] Duffie, D. and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance*, 6, 379-406.

[23] Duffie, D. and K.J. Singleton, 1997, An Econometric Model of the Term Structure of Interest Rates, *Journal of Finance*, 52, 1287-1321.

[24] Eichenbaum, M. and C.L.Evans, 1995, Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates, *Quarterly Journal of Economics*, 110:4, 975-1009..

[25] Eraker, B., 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *Journal of Finance*, 59, 1367-1403.

[26] Eraker, B., M. Johannes and N. Polson, 2003, The Impact of Jumps in Equity Index Volatility and Returns, *Journal of Finance*, 58, 1269-1300.

[27] Fruhwirth-Schnatter, Sylvia, 1994, Applied State-Space Modelling of non-Gaussian time-series using integration-based Kalman filtering, *Statistics and Computing*, 4, 259-269.

[28] Gallant, A.R. and G. Tauchen, 1998, Reprojecting Partially Observed Systems with Application to Interest Rate Diffusion, *Journal of the American Statistical Association*, 93, 10-24.

[29] Gallmeyer, Michael F., Burton Hollifield and Stanley E. Zin, 2005, Taylor Rules, McCullum Rules, and the Term Structure of Interest Rates, *Working Paper*.

[30] Hore, Satadru, Hedibert Lopes and Robert McCulloch, 2009, Put Options in General Equilibrium, *Working Paper*
http://www.rob-mcculloch.org/papers/option1.pdf

[31] Hordahl, P, O Tristani and D. Vestin, 2006, A Joint Econometric and Term Structure Dynamics, *Journal of Econometrics*, 131, 405-444..

[32] Jacquier, E., N.G. Polson and P.E. Rossi, 1994, Bayesian Analysis of Stochastic Volatility Models (with discussion), *Journal of Business and Economic Statistics*, 12, 371-417.

[33] Johannes, M. and N. Polson,2005, MCMC Methods for Financial Econometrics, *Working Paper*.

[34] Julier, S. and J. Uhlmann, 1997, A New Extension of the Kalman Filter to Non-Linear Systems, *Proceedings SPIE Signal Processing, Sensor Fusion, and Target Recognition VI*, 3068, 182-193.

[35] Kim, S., N. Shepherd and S. Chib, 1998, Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *Review of Economic Studies*, 65:3, 361-393.

[36] Koijen, Ralph and Jules H. van Binsbergen, 2008, Predictive Regressions: A Present Value Approach, *Working Paper*.

[37] Longstaff, F., 1989, A Non-Linear General Equilibrium Model of the Term Structure of Interest Rates, *Journal of Finance*, 23, 1259-1282.

[38] Mehra, Rajnish and Edward Prescott, 1985, The Equity Premium: A Puzzle, *Journal of Monetary Economics*, 15, 145-161.

[39] Piazzesi, M., 2005, Bond Yields and the Federal Reserve, *Journal of Political Economy*, 113:2, 311-344.

[40] Ravenna, Federico and Juha Seppala, 2007, Monetary Policy, Expected Inflation and Inflation Risk-Premia, *Working Paper*.

[41] Sargent, Thomas J., 1979, A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure, *Journal of Monetary Economics*, 35, 245-74.

[42] Sims, C.A., 2001, Comments on Cogley and Sargent's "Evolving Post-World War II U.S. Inflation Dynamics", In B.S. Bernanke and K. Rogoff. (eds) *NBER Macroeconomics Annual 2001*, 16, 373-279, MIT Press.

[43] Stock, J.H., 2001, Comments on Cogley and Sargent's "Evolving Post-World War II U.S. Inflation Dynamics", In B.S. Bernanke and K. Rogoff. (eds) *NBER Macroeconomics Annual 2001*, 16, 373-279, MIT Press.

[44] Weil, P., 1989, The equity-premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics*, 24(3), 401-421.

[45] Yared, F. and P. Veronesi, 2000, Short and Long Horizon Term and Inflation Risk Premia in the US Term Structure: Evidence from an Integrated Model for Nominal and Real Bond Prices under Regime Shifts, *Working Paper*, The University of Chicago.

Table I: Posterior medians of parameter estimates along with the 2.5-97.5 quantile of each parameter distribution that are obtained from the markov chain generated from the MCMC sampling. Each Gibbs sampler is run for 25000 iterations and the first 15000 draws are discarded. The estimates presented here are from the last 10000 draws. "No Bonds" represents parameters estimates without the yield curve. "With Bonds" respresents parameters estimated with the yield data of 1-5 years. Parameters are reported in monthly quantities. The **tight** model is with pricing error (ϵ_τ) of 3bp, the **loose** prior with 20bp and the **joint** prior with 10 basis points.

	No Bonds	*With Bonds*		
		Tight	Loose	Joint
α_0^c	0.0025 (0.0027,0.0044)	0.0068 (0.0067,0.0070)	0.0035 (0.0032,0.0037)	0.0016 (0.0014,0.0019)
$1-\alpha_1^c$	-0.4863 (-0.6521,-0.2736)	-0.5131 (-0.5157,-0.5100)	-0.4711 (-0.4749,-0.4691)	-0.5190 (-0.5247,-0.5074)
α_2^c	-0.2551 (-0.4646,-0.0602)	-0.3877 (-0.3881,-0.3869)	-0.2487 (-0.2503,-0.2471)	-0.3167 (-0.3573,-0.2867)
α_0^q	0.0005 (-0.0002,0.0017)	0.0005 (0.0003,0.0007)	0.0008 (0.0007,0.0010)	0.0004 (0.0003,0.0005)
α_1^q	-0.1019 (-0.5610,0.1577)	-0.3021 (-0.3048,-0.2992)	-0.0986 (-0.1024,-0.0925)	-0.3409 (-0.3028,-0.3392)
$1-\alpha_2^q$	0.9009 (0.7825,0.9705)	0.9528 (0.9311,0.9612)	0.9025 (0.8819,0.9332)	0.9613 (0.9602,0.9625)
β_0^c	0.0004 (0.0002,0.0005)	0.0033 (0.0028,0.0035)	-0.0024 (-0.0041,-0.0011)	-0.0002 (0.0005,0.0013)
$1-\beta_1^c$	0.8551 (0.8228,0.8867)	0.8635 (0.8430,0.8838)	0.8597 (0.8583,0.8616)	0.8602 (0.7974,0.8805)
β_0^q	0.0002 (0.0001,0.0003)	-0.0002 (-0.0002,-0.0001)	0.0008 (0.0007,0.0010)	0.0003 (0.0002,0.0004)
$1-\beta_1^q$	0.8567 (0.8244,0.8893)	0.8576 (0.8376,0.8880)	0.8568 (0.8549,0.8574)	0.8588 (0.8291,0.8647)
ϵ^c	0.0014 (0.0013,0.0016)	0.0019 (0.0018,0.0021)	0.0013 (0.0012,0.0013)	0.0010 (0.0009,0.0012)
ϵ^q	0.0009 (0.0008,0.0010)	0.0019 (0.0017,0.0020)	0.0015 (0.0014,0.0016)	0.0012 (0.0009,0.0019)
σ^c	0.0033 (0.0028,0.0035)	0.0159 (0.0151,0.0165)	0.0013 (0.0012,0.0015)	0.0088 (0.0082,0.0090)
σ^q	0.0028 (0.0026,0.0032)	0.0767 (0.0733,0.0800)	0.0006 (0.0006,0.0006)	0.0100 (0.0065,0.0119)
ρ_{cq}	-0.0174 (-0.0356,-0.0047)	0.5378 (0.5500,0.3414)	-0.1491 (-0.1807,-0.1177)	-0.0024 (-0.0324,0.0266)
ρ_1	0.0331 (-0.024,0.1161)	-0.6760 (-0.6942,-0.6699)	0.0487 (-0.0280,0.1284)	-0.5454 (-0.5470,-0.5442)
ρ_2	0.0502 (-0.046,0.1506)	0.4993 (0.4564,0.5294)	0.2001 (0.0201,0.3251)	0.5409 (0.5401,0.5422)
γ	- -	4.7080 (4.6504,4.8041)	0.4534 (0.2241,0.6341)	3.7442 (3.5897,3.8183)
ϕ	- -	0.0070 (0.0100,0.0188)	0.0060 (0.0010,0.0100)	0.0170 (0.0150,0.0198)

Table II: Annual estimates of loadings on the risk-premia in (15). Given the distribution of parameter values in Table 1, the loading on the risk premia for ϑ_t^c ($\rho_1\gamma\sigma_c B(\tau)\cdot 12$) and ϑ_t^q ($\rho_2\sigma_q C(\tau)\cdot 12$) are computed for each bond of maturity τ. The median level, as well as the 2.5-97.5 quantiles, of the loadings are reported. The **tight** model is with pricing error (ϵ_τ) of 3bp, the **loose** prior with 20bp and the **joint** prior with 10 basis points.

τ	**tight**		**loose**		**joint**	
	ϑ^c	ϑ^q	ϑ^c	ϑ^q	ϑ^c	ϑ^q
1	2.0051 (1.7964,2.2394)	0.7713 (0.5656,0.8852)	0.0427 (-0.0269,0.1422)	-0.0107 (-0.0178,-0.0011)	0.3809 (0.3090,0.4144)	-0.2055 (-0.2351,-0.1914)
2	2.4045 (2.1223,2.7071)	2.3858 (1.9470,2.5698)	0.1059 (-0.0589,0.2893)	-0.0147 (-0.0245,-0.0016)	0.1534 (0.0717,0.1991)	-0.4709 (-0.5316,-0.4354)
3	2.9770 (2.6162,3.3532)	4.6975 (4.0535,4.9073)	0.1280 (-0.0710,0.3514)	-0.0162 (-0.0271,-0.0017)	-0.1815 (-0.2638,-0.1215)	-0.8651 (-0.9531,-0.8017)
4	3.7967 (3.3671,4.2470)	7.9906 (7.2593,8.3012)	0.1363 (-0.0762,0.3761)	-0.0168 (-0.0281,-0.0018)	-0.6769 (-0.8097,-0.5978)	-1.4484 (-1.5717,-1.3479)
5	4.9768 (4.5015,5.4803)	12.6864 (12.1079,13.3425)	0.1395 (-0.0782,0.3863)	-0.0170 (-0.0284,-0.0018)	-1.3953 (-1.7174,-1.2539)	-2.3060 (-2.5977,-2.1426)

Table III: This table reports the posterior median and the 2.5-97.5 quantile of pricing error in basis points on monthly yields. The **tight** model is with pricing error (ϵ_τ) of 3bp, the **loose** prior with 20bp and the **joint** prior with 10 basis points.

τ	**tight**	**loose**	**joint**
1	1.5723 (0.9856,1.9250)	16.2781 (12.1895,19.2196)	7.5793 (5.5662,9.9867)
2	2.8993 (1.4287,3.6529)	15.3845 (10.1819,21.6840)	7.2016 (6.9078,9.9985)
3	2.2774 (1.1876,3.8901)	15.7334 (11.1671,23.1775)	5.6390 (3.470,8.8119)
4	2.8703 (1.2432,3.5490)	14.1662 (9.8751,22.1446)	5.3388 (4.2179,8.8201)
5	1.9486 (1.0382,3.6674)	14.1819 (8.8653,21.3884)	3.4455 (2.1890,8.8659)

Table IV: This table reports correlation of bond return and risk-premia from each prior setting. First, we construct a series of yearly bond returns, by taking an n year bond at time t and sell it the following year as an $n-1$ year bond. Our sample in this case is Fama-Bliss 2-5 year yields from 1959-2005. Then, we construct the annualized risk-premia from (15) - $(\rho_1 \gamma \sigma_c B(\tau) \vartheta_t^c + \rho_2 \sigma_q C(\tau) \vartheta_t^q) \cdot 12$ by taking the posterior distribution of parameters and state variables from each setting. We compute correlation of each time-series in the posterior distribution with the bond return computed from the sample. The 2.5-97.5 quantile of the distribution of correlations, along with the median, is reported for each prior setting. The **tight** model is with pricing error (ϵ_τ) of 3bp, the **loose** prior with 20bp and the **joint** prior with 10 basis points.

τ	**tight**	**loose**	**joint**
2	-0.6090 (-0.6491,-0.5772)	0.0316 (-0.1080,0.1587)	0.1901 (0.0655,0.2861)
3	-0.5802 (-0.6345,-0.5392)	0.0294 (-0.1082,0.1569)	0.2201 (0.0844,0.3013)
4	-0.5730 (-0.6330,-0.5290)	0.0289 (-0.1069,0.1560)	0.2225 (0.0788,0.3100)
5	-0.5606 (-0.6225,-0.5154)	0.0300 (-0.1061,0.1561)	0.2191 (0.0711,0.3100)

Table V: In this Table we present the posterior median and 2.5-97.5 quantile of the distribution of R^2's from predictability regression. First, we construct a series of yearly bond returns, by taking an n year bond at time t and sell it the following year as an $n-1$ year bond. We compute excess return by subtracting off the yield on the 1-year bond. Our sample in this case is Fama-Bliss 2-5 year yields from 1959-2005. Then, we construct the annualized risk-premia from (15) - $(\rho_1 \gamma \sigma_c B(\tau) \vartheta_t^c + \rho_2 \sigma_q C(\tau) \vartheta_t^q) \cdot 12$ by taking the posterior distribution of parameters and state variables from the **joint** setting. Then we regress the bond return at year $t+1$ on the risk-premia computed at year t and report the R^2's from each setting. For comparison purposes, we also provide the R^2's from ADP(2007) who also look at Fama-Bliss bonds in a similar time-period in quarterly frequency.

τ	**joint**	ADP(2007)
2	0.0361 (0.0046,0.0818)	-
3	0.0484 (0.0075,0.0908)	0.096
4	0.0495 (0.0064,0.0961)	-
5	0.0480 (0.0055,0.0961)	0.096

Table VI: This table presents the predictability coefficients of regressing excess bond returns on ϑ_t^c and ϑ_t^q that were filtered under **joint** and **tight** setting. First, we construct a series of yearly bond returns, by taking an n year bond at time t and sell it the following year as an $n-1$ year bond. We compute excess return by subtracting off the yield on the 1-year bond. Our sample in this case is Fama-Bliss 2-5 year yields from 1959-2005. Taking the full distribution of the posterior distribution of filtered state variables ϑ_t^c and ϑ_t^q under each prior setting, we regress excess bond return on both of these filtered series. We report the median and the 2.5-97.5 quantile of the regression coefficient from the regression

$$r_{t+1} = a_0 + a_1\vartheta_t^c + a_2\vartheta_t^q + \epsilon$$

Then, we present the median and 2.5-97.5 quantile of the distribution of the corresponding coefficients from the risk-premia relationship $(\rho_1\gamma\sigma_c B(\tau)\vartheta_t^c + \rho_2\sigma_q C(\tau)\vartheta_t^q) \cdot 12$, which is repeated from Table 2. The top panel presents the **joint** and the bottom panel presents the **tight** setting.

	joint			
τ	a_1	$\rho_1\gamma\sigma_c B(\tau)$	a_2	$\rho_2\sigma_q C(\tau)$
2	0.5925 (-1.3951,2.6179)	0.1534 (0.0717,0.1991)	-1.4057 (-2.4483,-0.2808)	-0.4709 (-0.5316,-0.4354)
3	1.1318 (-2.4887,4.8042)	-0.1815 (-0.2638,-0.1215)	-3.0606 (-4.8747,-1.0011)	-0.8651 (-0.9531,-0.8017)
4	1.6537 (-3.3271,6.5707)	-0.6769 (-0.8097,-0.5978)	-4.4933 (-6.9545,-1.6563)	-1.4484 (-1.5717,-1.3479)
5	2.1126 (-4.0281,8.0692)	-1.3953 (-1.7174,-1.2539)	-5.6178 (-8.5989,-2.2256)	-2.3060 (-2.5977,-2.1426)
	tight			
2	-0.5409 (-0.6392,-0.4208)	2.4045 (2.1233,2.7071)	-0.7145 (-0.9940,-0.5107)	2.3858 (1.9470,2.5698)
3	-0.8175 (-1.0029,-0.6055)	2.9770 (2.6162,3.3532)	-1.4603 (-1.9631,1.0718)	4.6975 (4.0535,4.9073)
4	-1.0384 (-1.3032,-0.7419)	3.7967 (3.3671,4.2470)	-2.1430 (-2.7936,-1.6036)	7.9906 (7.2593,8.3012)
5	-1.2283 (-1.5535,-0.8627)	4.9768 (4.5015,5.4803)	-2.5945 (-3.3715,-1.9311)	12.6864 (12.1079,13.3425)

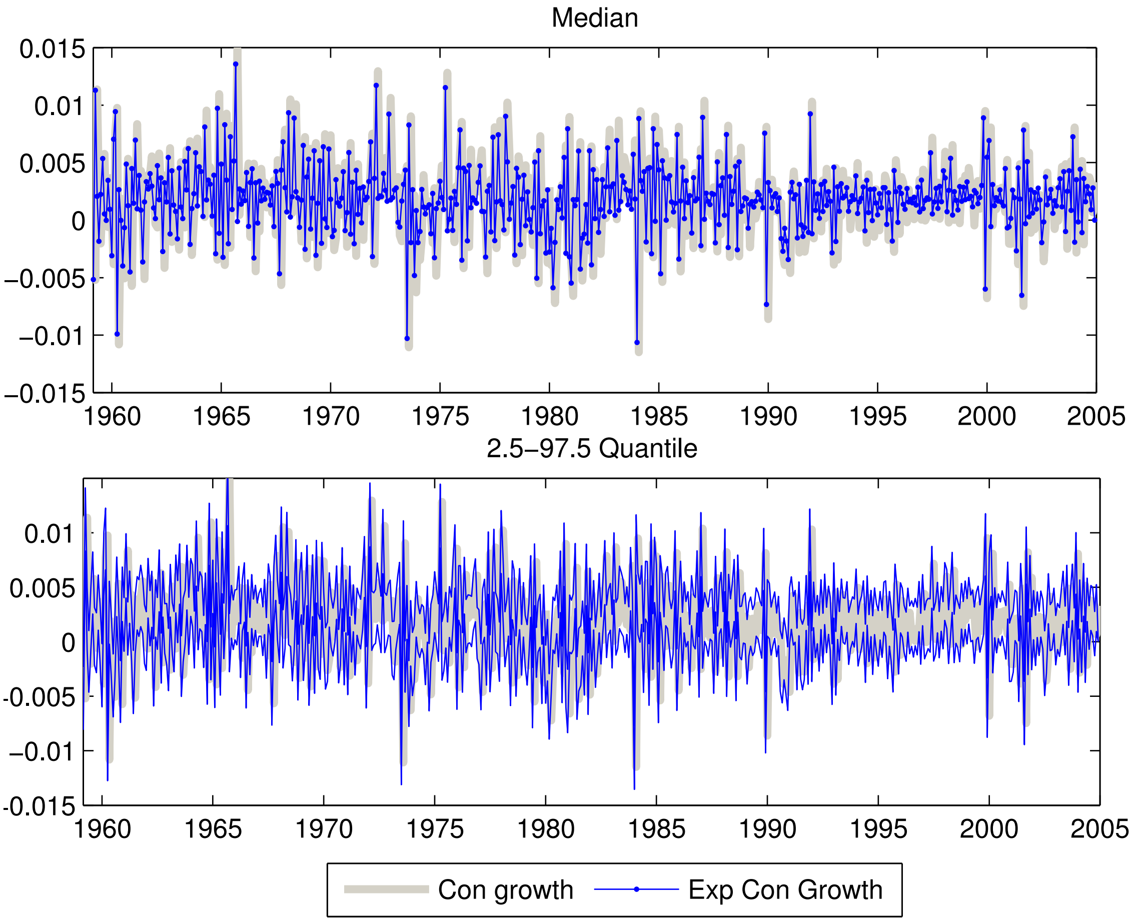

Figure 1: *Filtered and smoothed expected consumption growth Rate from the* ***loose*** *specification. The top panel is the median of* μ_t^c *and the bottom is the 2.5-97.5 quantile. The heavy background is data of monthly consumption growth.*

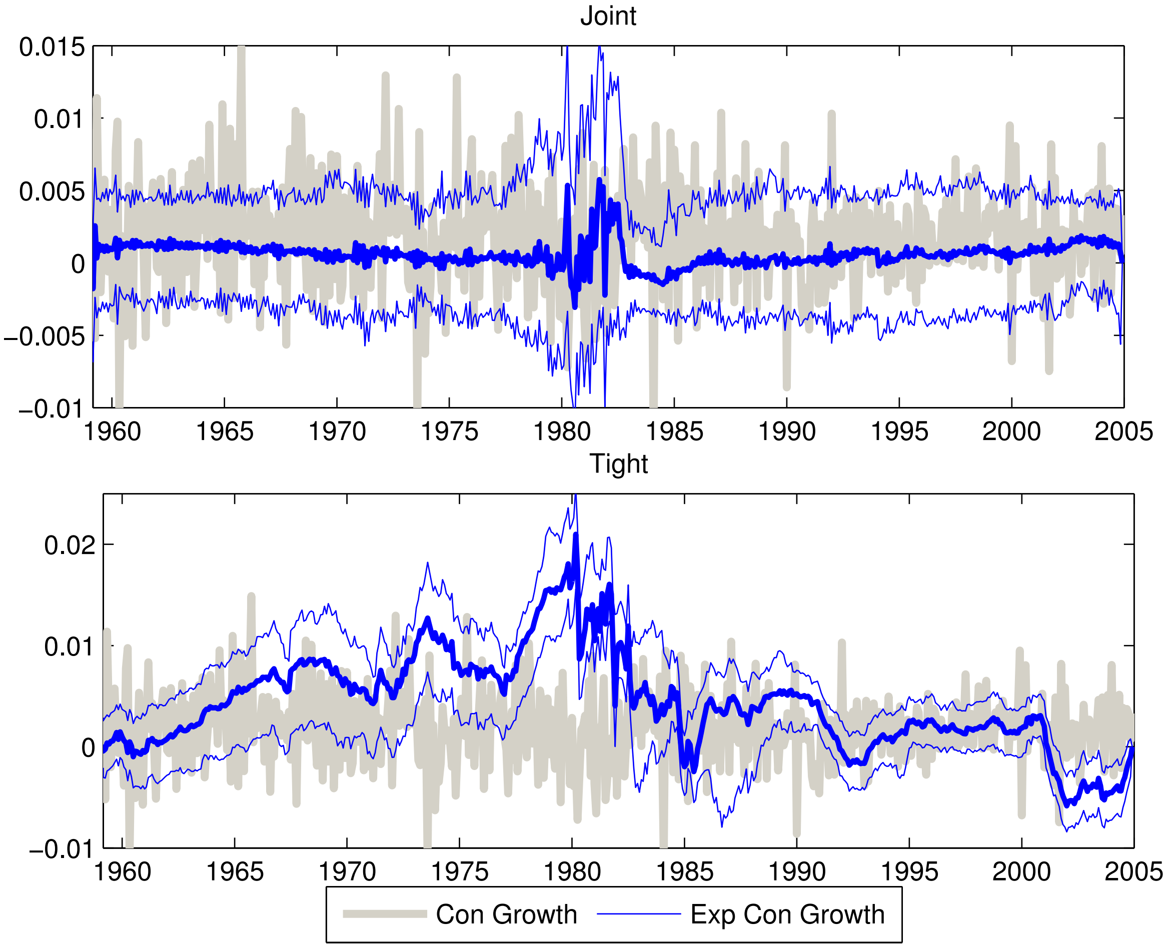

Figure 2: *Filtered and smoothed expected consumption growth Rate from the **joint** and **tight** specification. The top (bottom) panel reports the 2.5-median-97.5 quantile of the posterior distribution of* μ_t^c *from the **joint** (**tight**) prior setting. The heavy background is data of monthly consumption growth.*

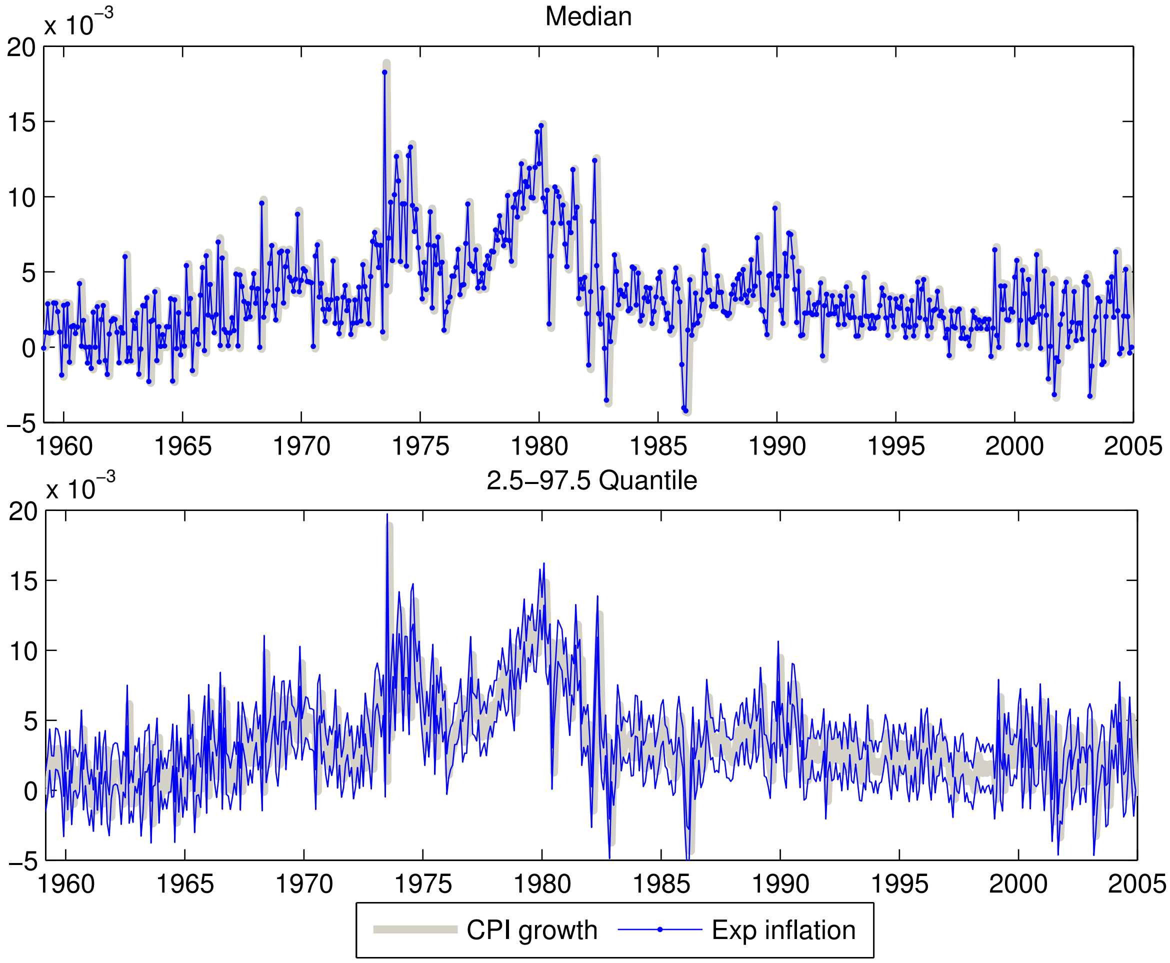

Figure 3: *Filtered and smoothed expected inflation from the **loose** specification. The top panel is the median of* μ_t^q *and the bottom is the 2.5-97.5 quantile. The heavy background is data of monthly CPI growth.*

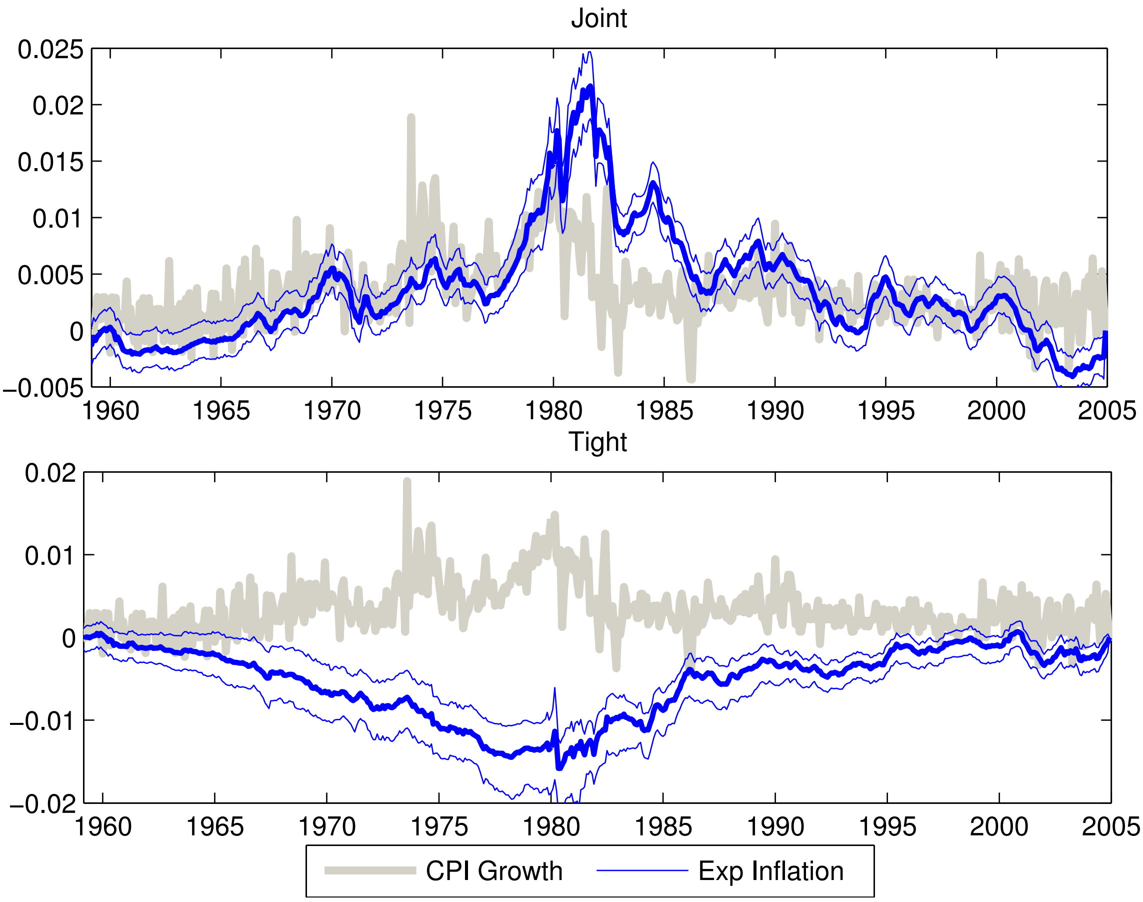

Figure 4: *Filtered and smoothed expected inflation Rate from the* ***joint*** *and* ***tight*** *specification. The top (bottom) panel reports the 2.5-median-97.5 quantile of the posterior distribution of* μ_t^q *from the* ***joint*** *(****tight*** *) prior setting. The heavy background is data of monthly CPI growth.*

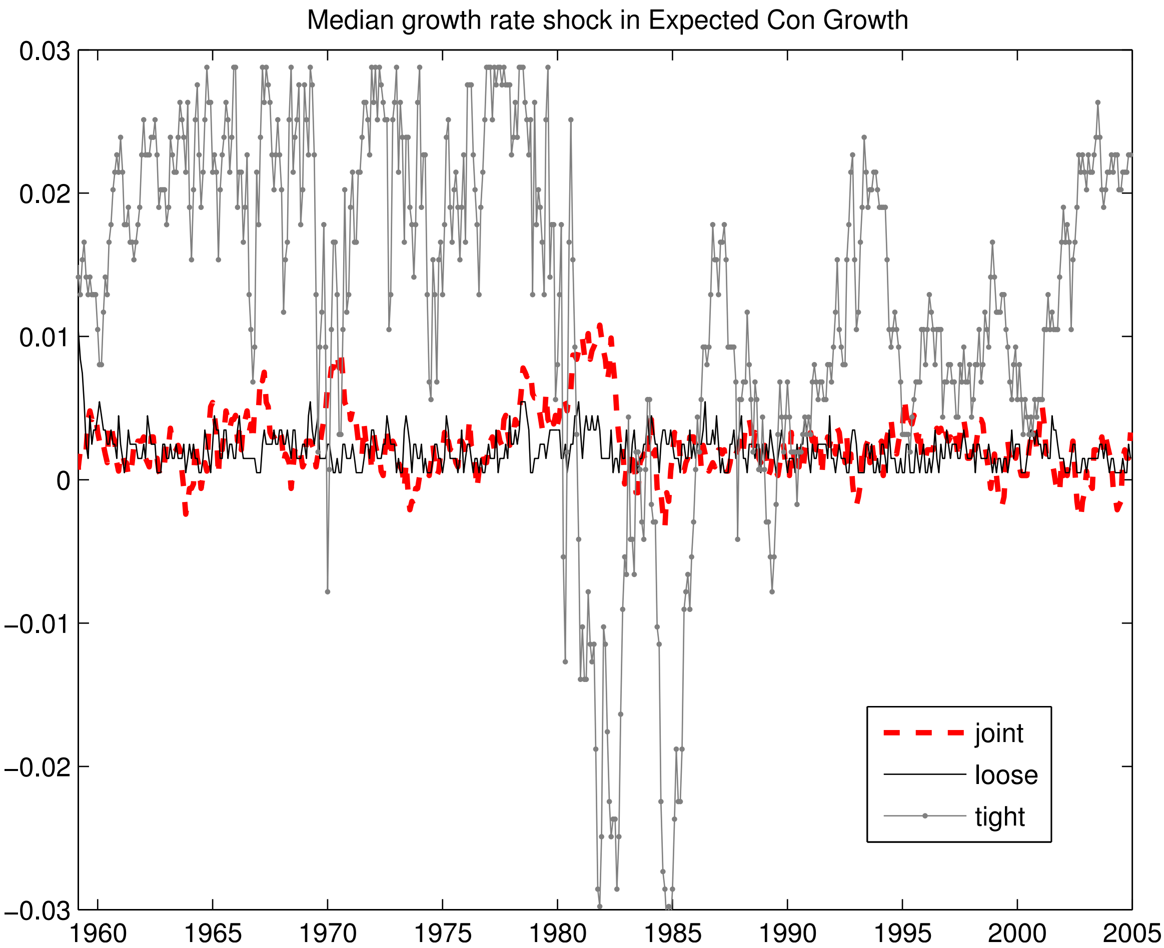

Figure 5: *Posterior median of the filtered and smoothed conditional volatility of expected consumption growth* ϑ_t^c*from the* ***tight*** *,**loose** and **joint** setting.*

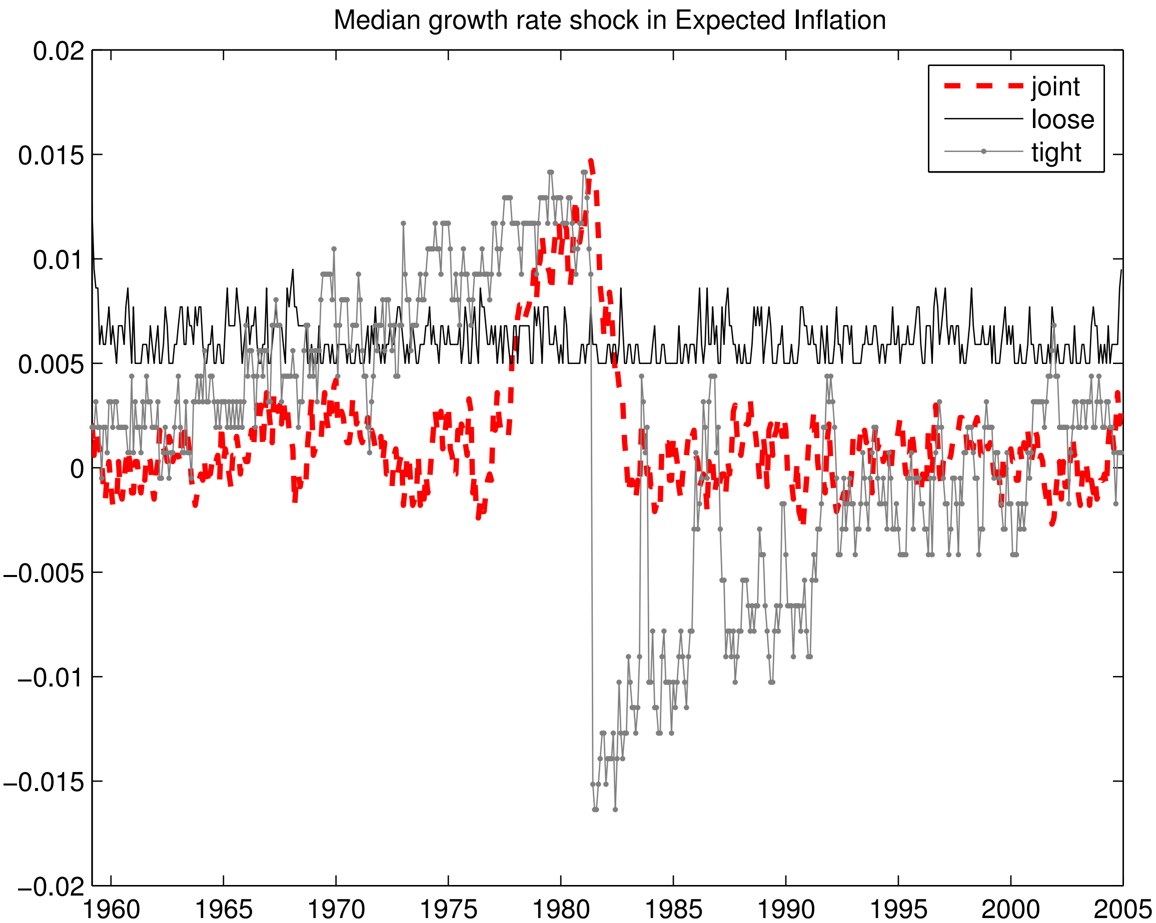

Figure 6: *Posterior median of the filtered and smoothed conditional volatility of expected inflation* ϑ_t^q*from the* ***tight*** *,**loose** and **joint** setting.*

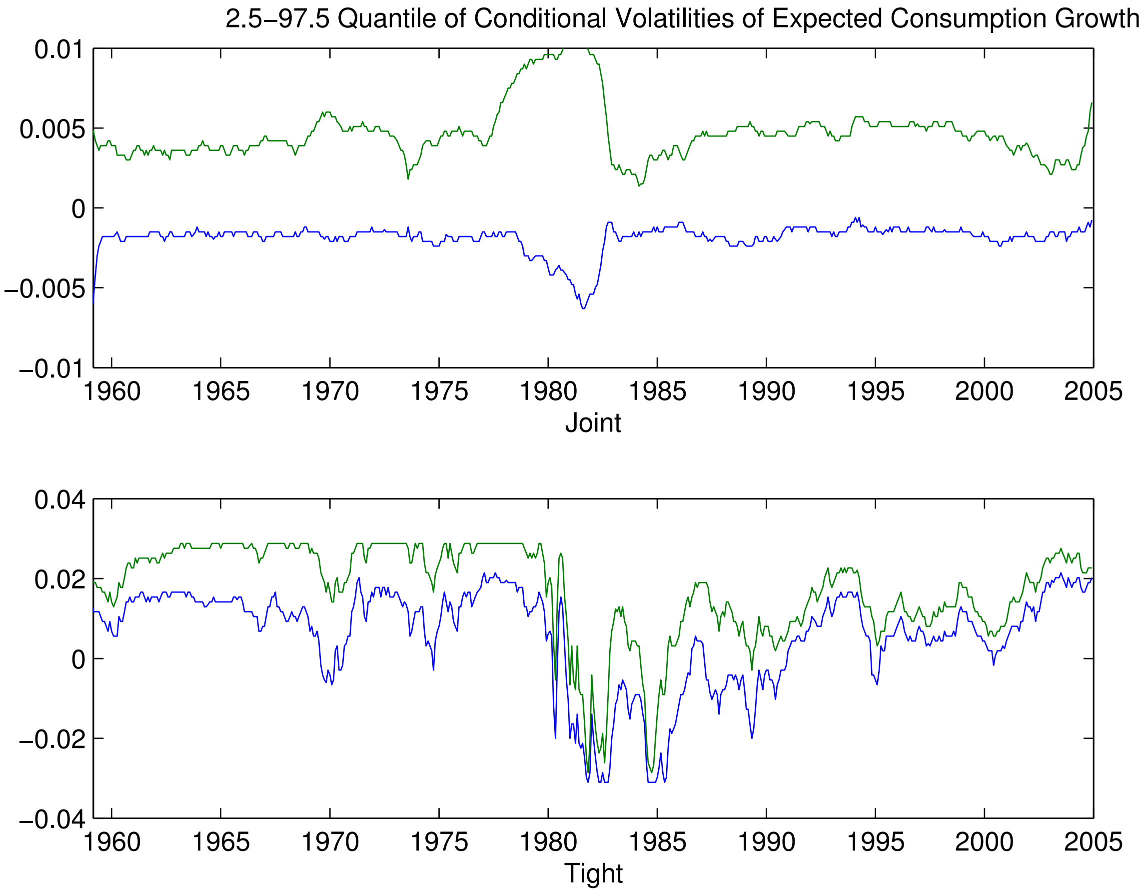

Figure 7: *The 2.5-97.5 Quantile of the filtered and smoothed conditional volatility of expected consumption growth.*

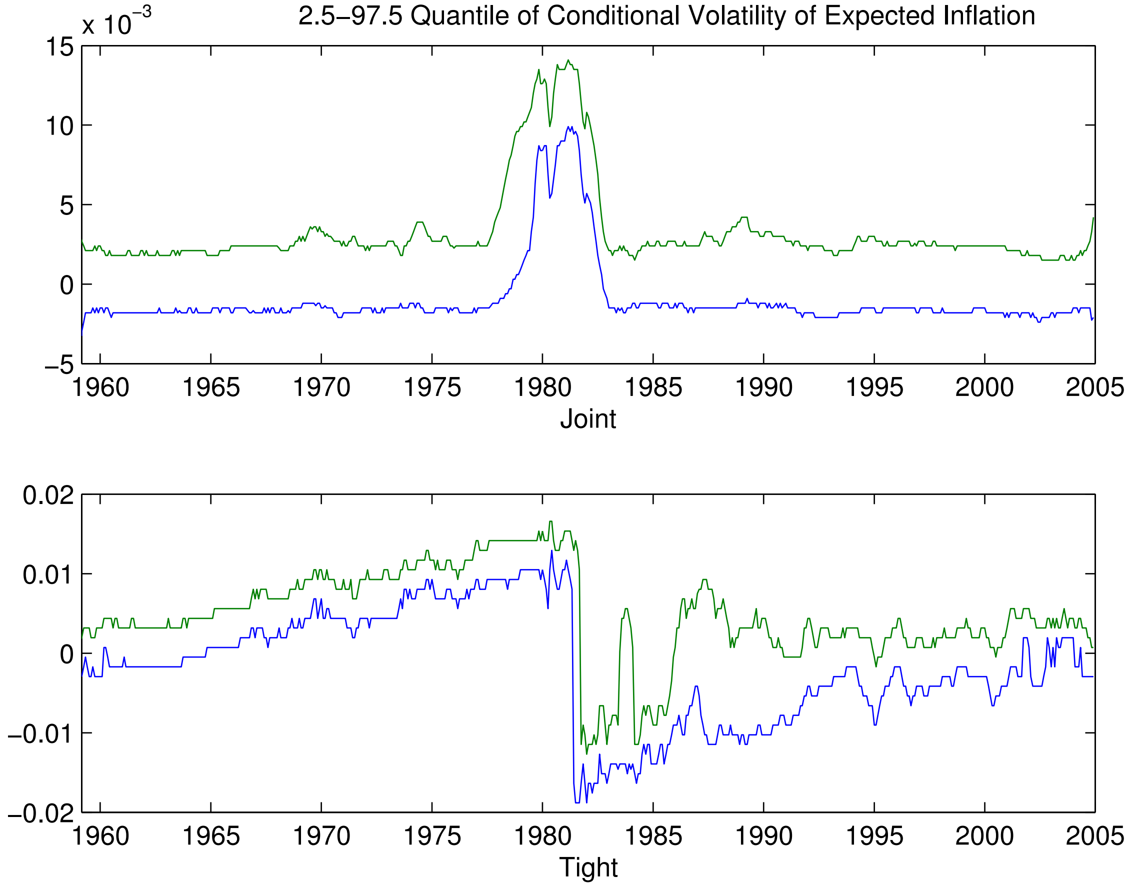

Figure 8: *The 2.5-97.5 Quantile of the filtered and smoothed conditional volatility of expected inflation.*

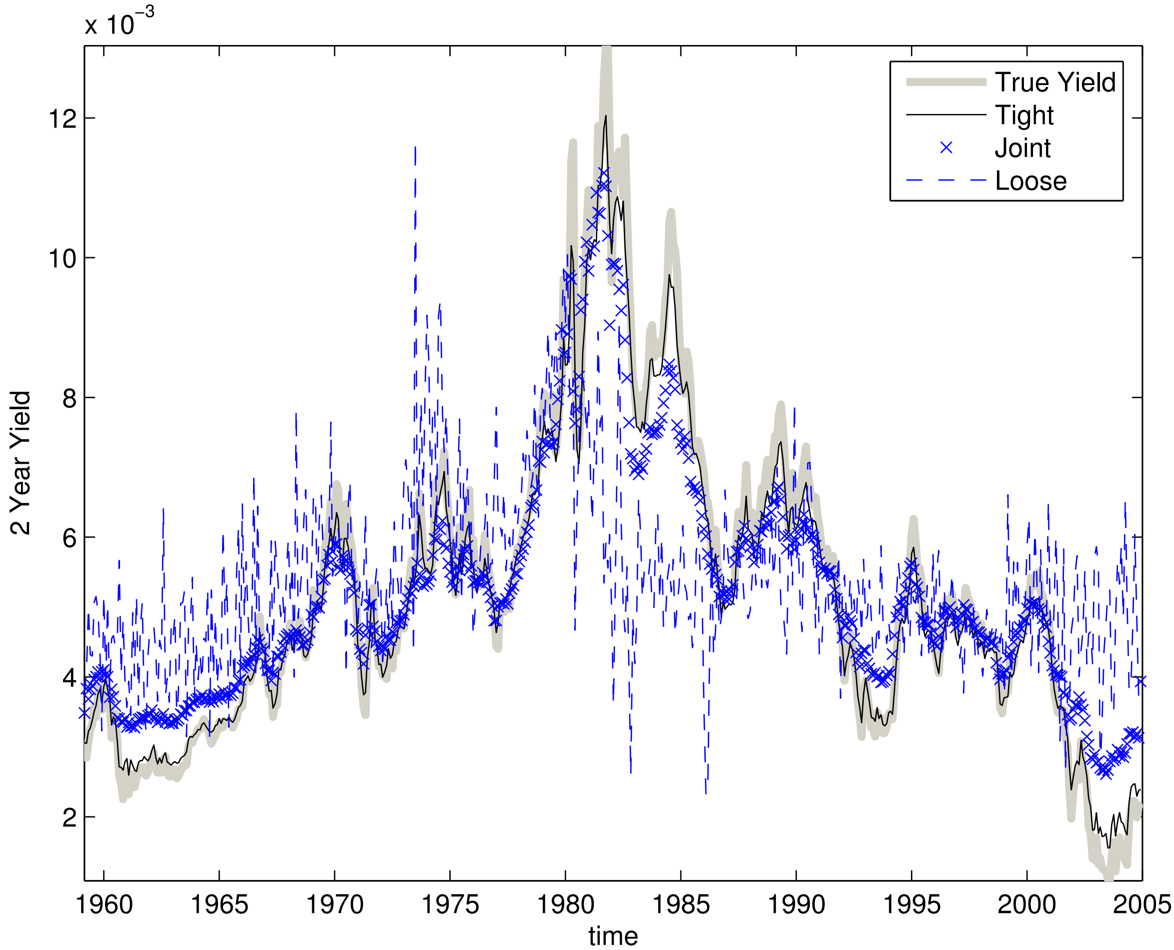

Figure 9: *Posterior median time-series estimate of the 2-year yield curve under the three prior settings.*

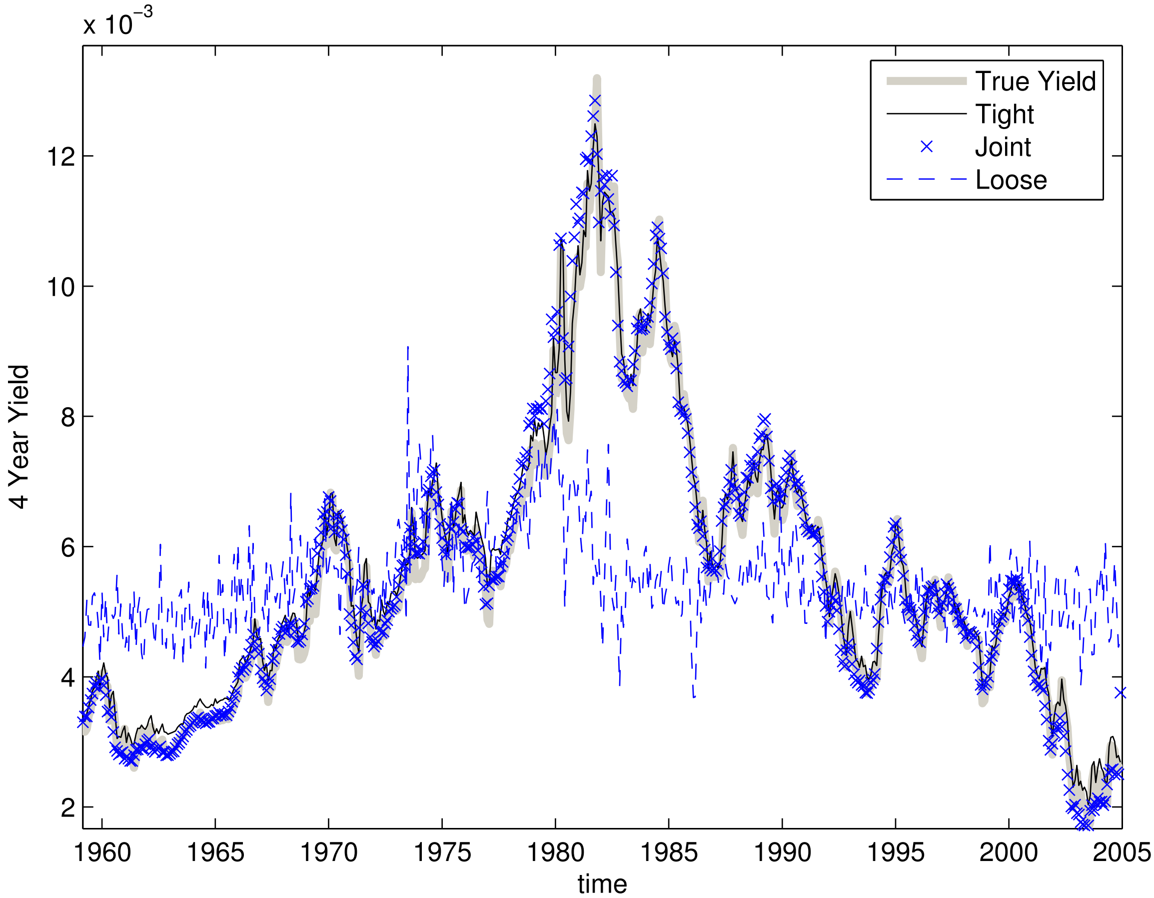

Figure 10: *Posterior Median of the 4-year yield curve under the three settings.*

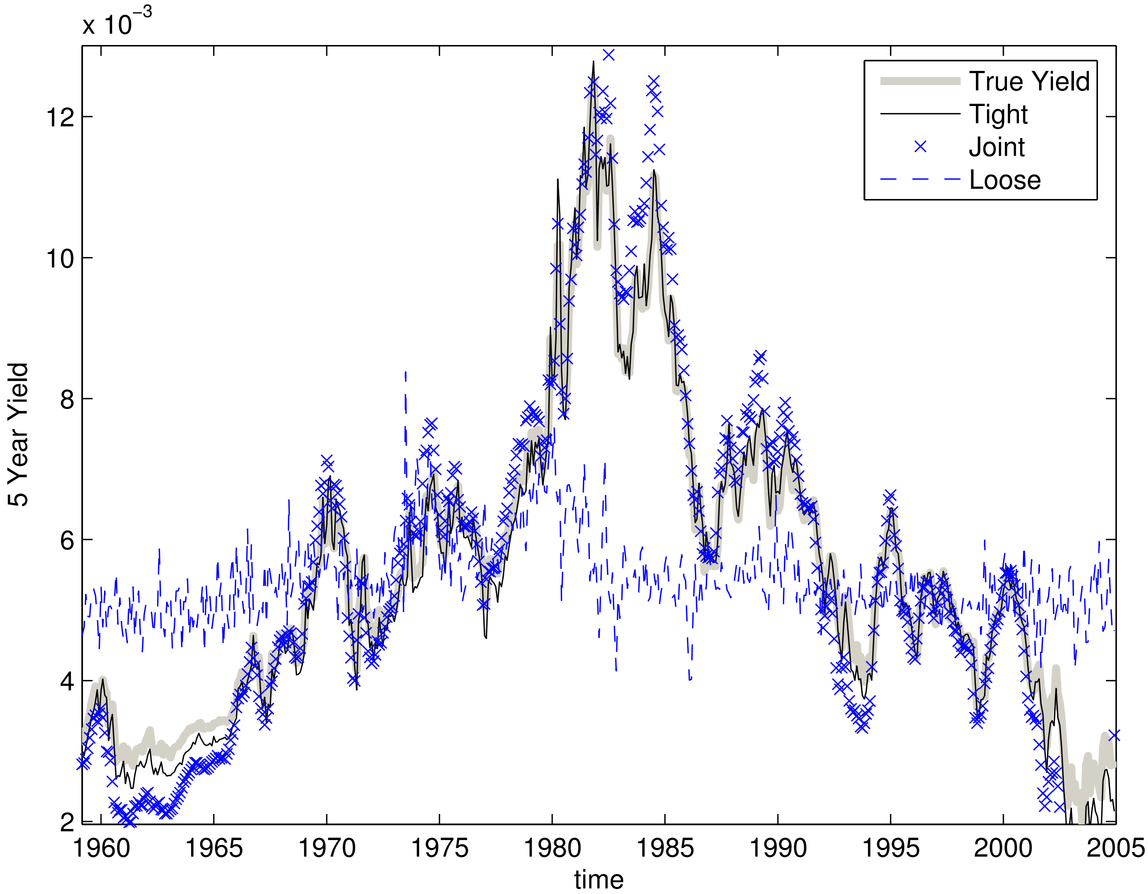

Figure 11: *Posterior Median of the 5-year yield curve under the three settings.*

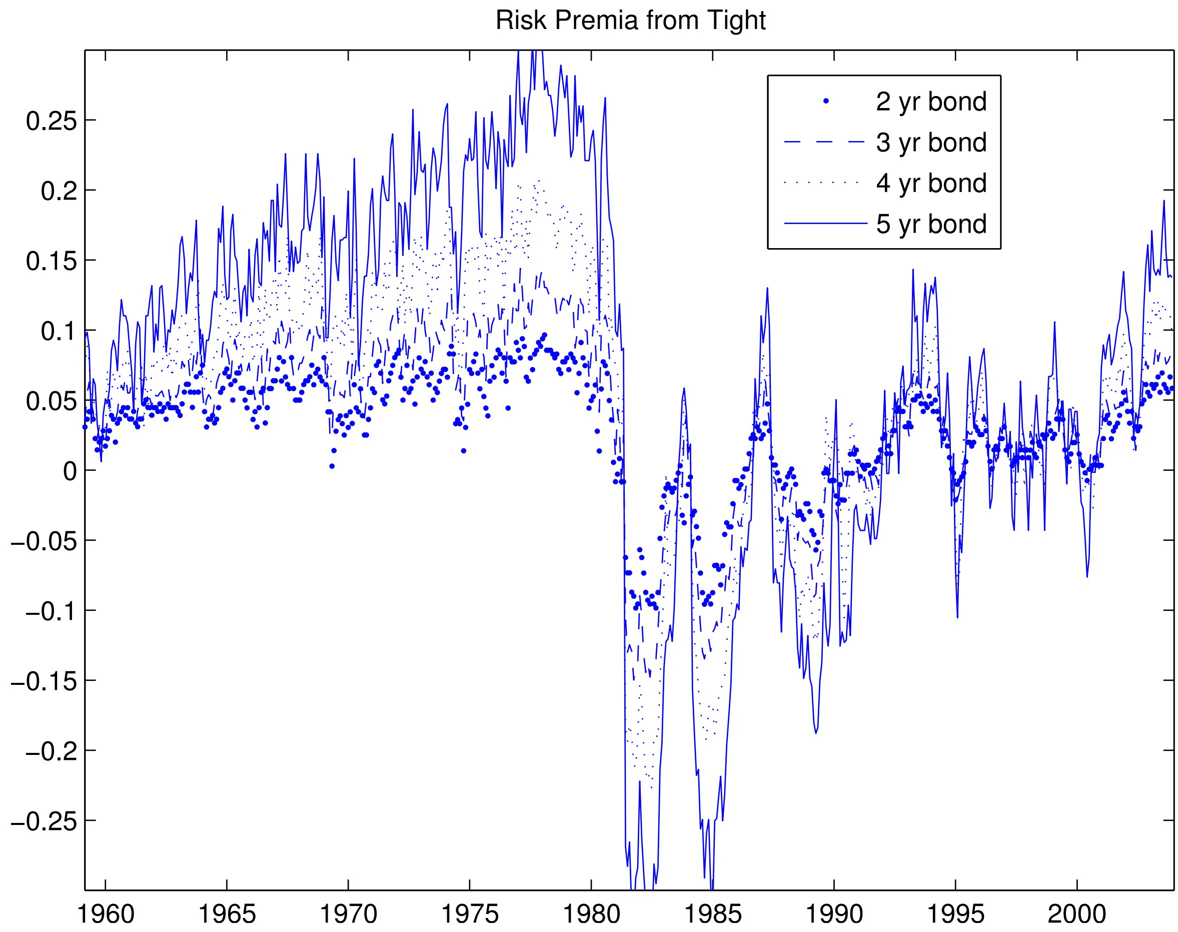

Figure 12: *Posterior Median of expected return on bonds from the* ***tight*** *setting. Taking parameter estimates and filtered state variables under* ***tight*** *setting, we form the posterior distribution of the annual average return of 5-year bonds according to (15) and report the median.*

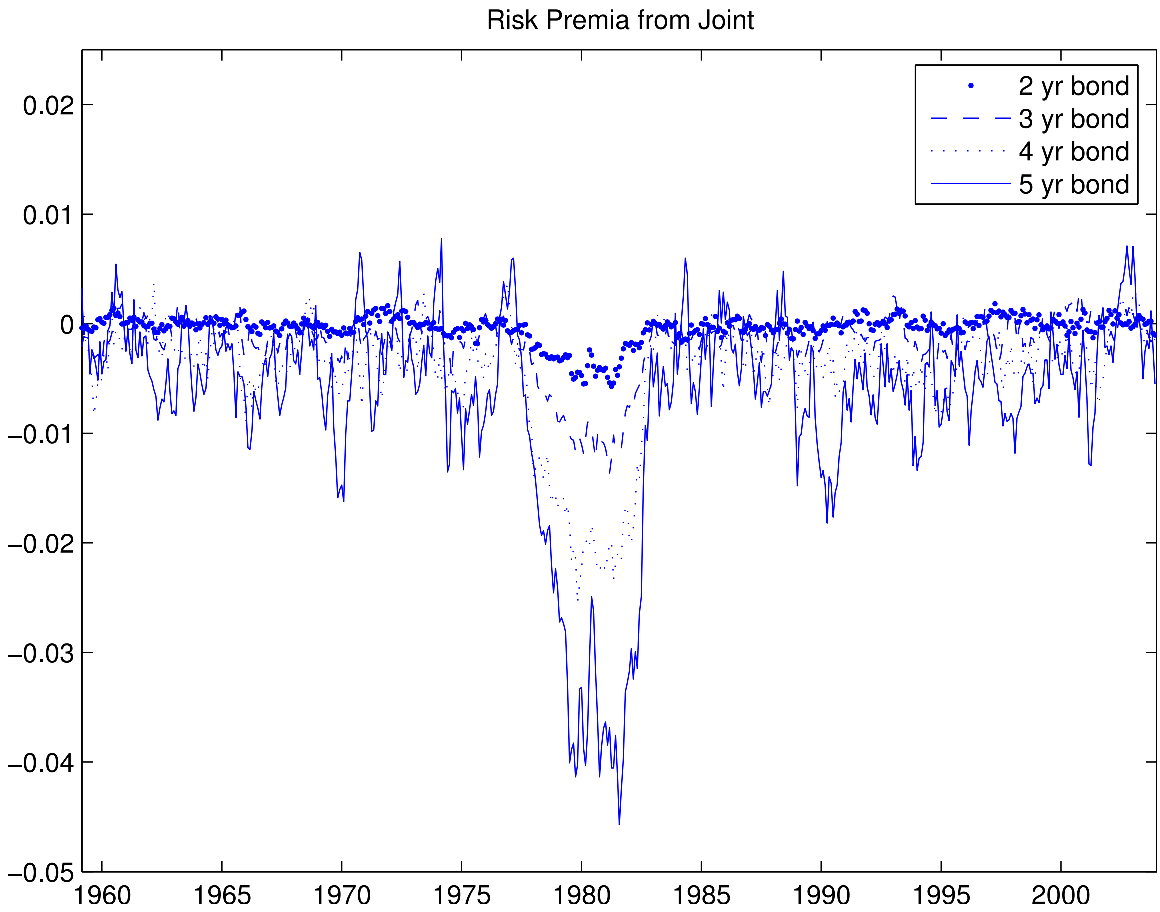

Figure 13: *Posterior Median of expected return on bonds from the* ***joint*** *setting. Taking parameter estimates and filtered state variables under* ***joint*** *setting, we form the posterior distribution of the annual average return of 5-year bonds according to (15) and report the median.*

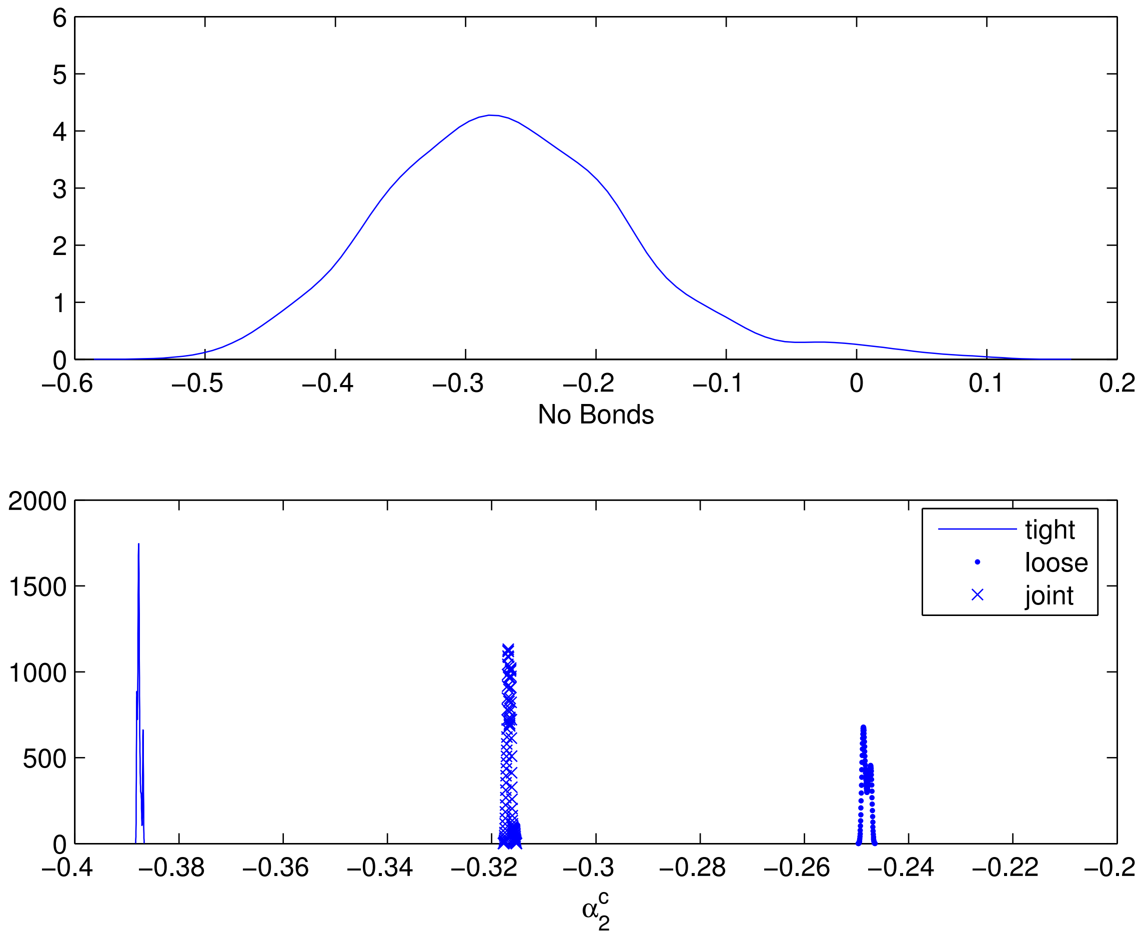

Figure 14: *Posterior distribution of* α_2^c *used to test the Fisher Hypothesis.*

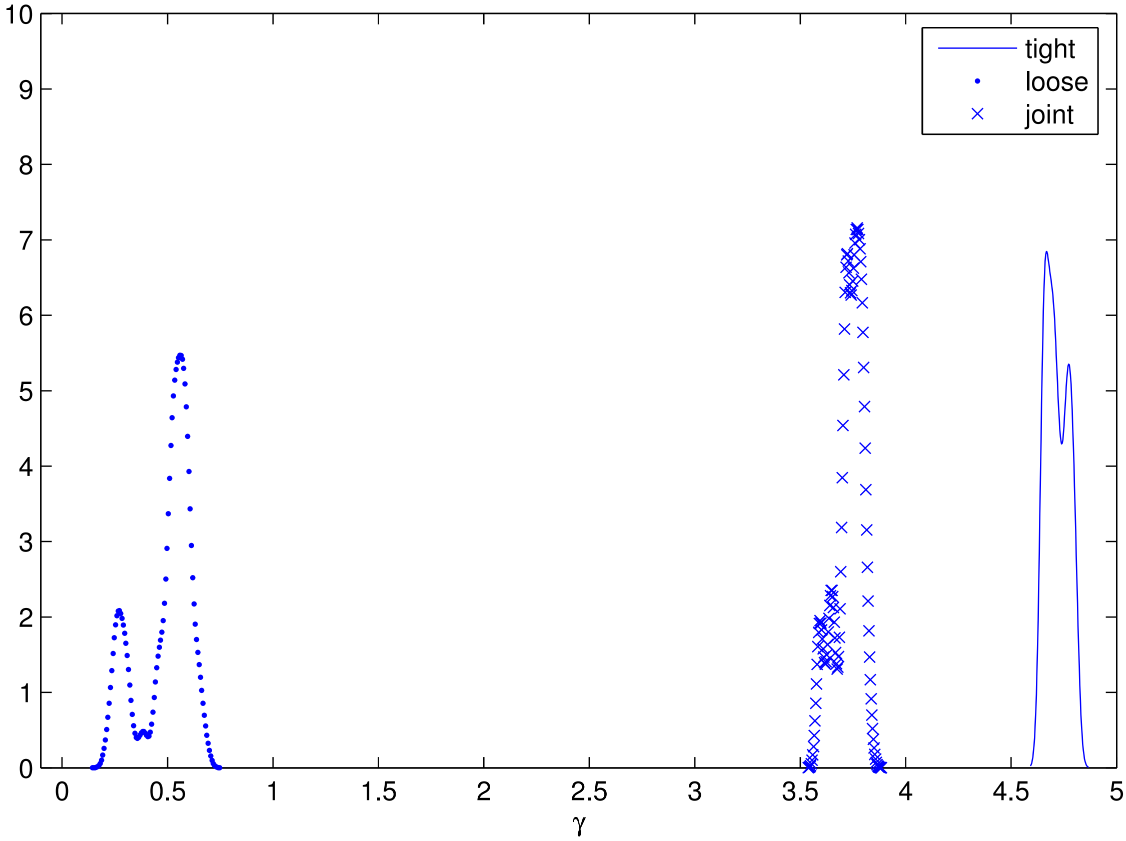

Figure 15: *Posterior distribution of the risk-aversion parameter, γ.*

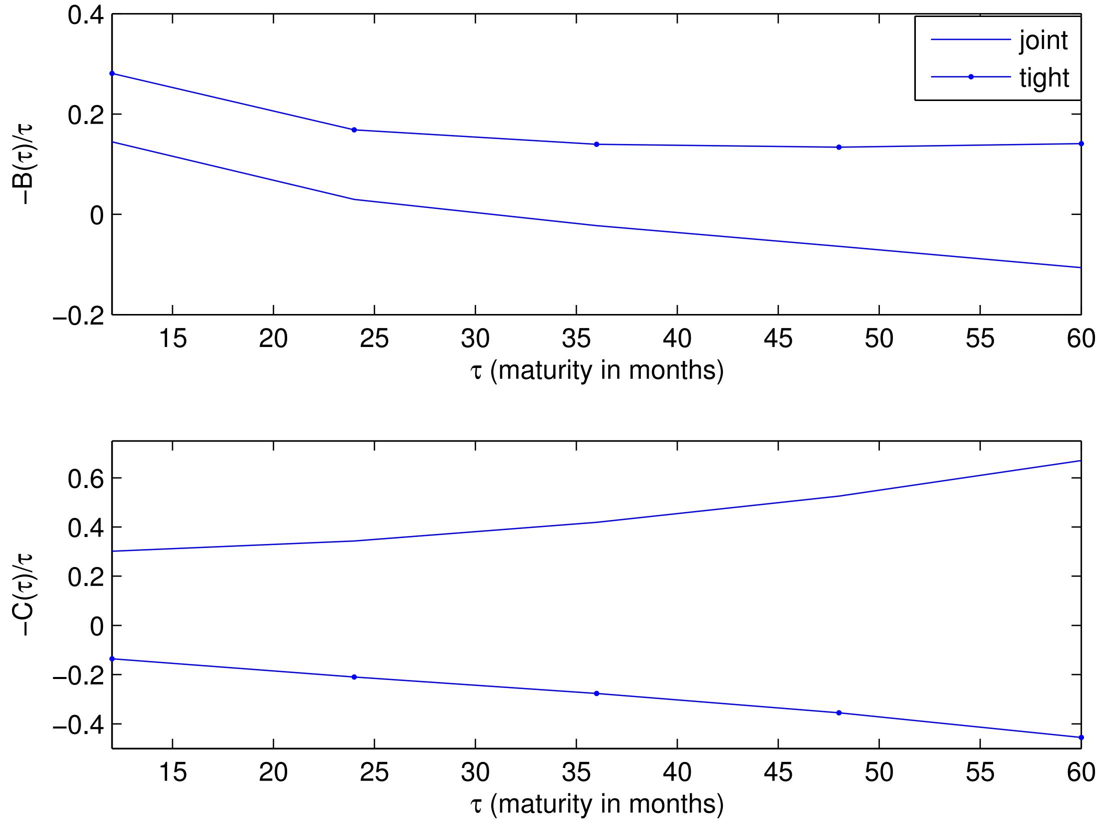

Figure 16: *Loading on expected consumption growth (top) and expected inflation (bottom) in the yield curve formula (14). Taking posterior estimates from* ***tight*** *and* ***joint*** *setting, we form the posterior distribution of* $-B(\tau)/\tau$ *and* $-C(\tau)/\tau$ *and report the median for each prior setting.*

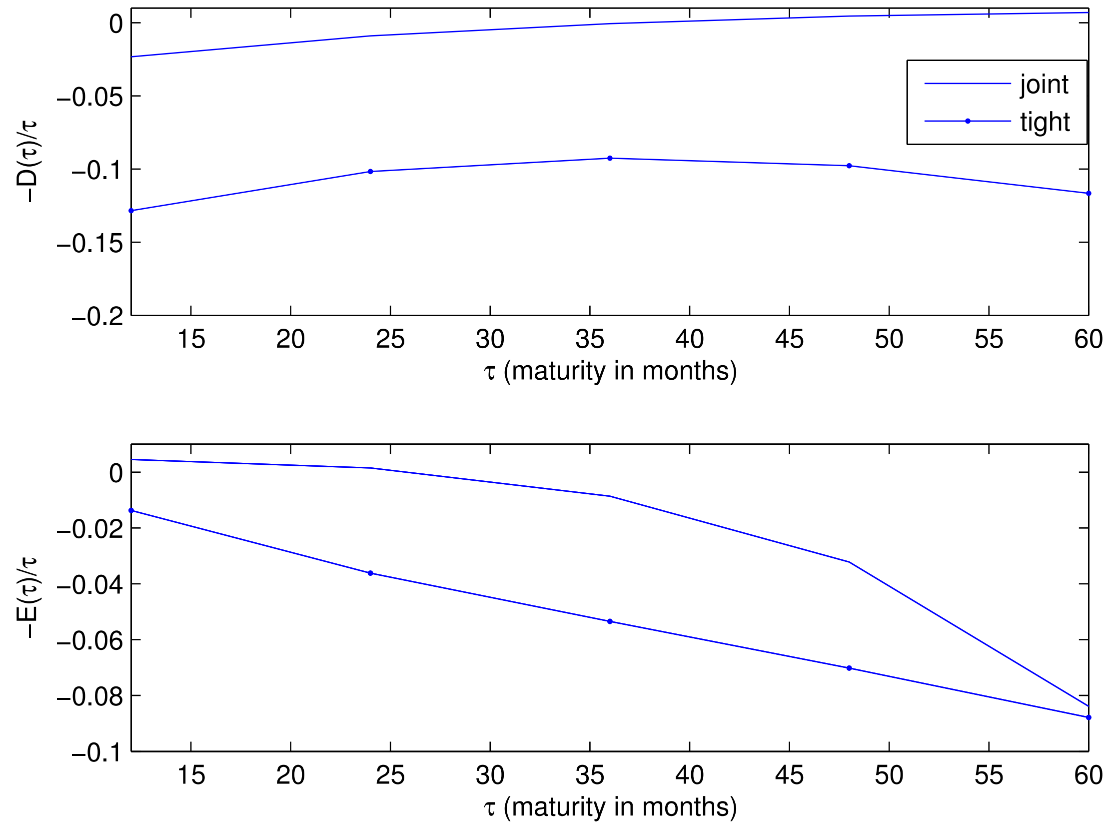

Figure 17: *Loading on conditional volatility of expected consumption growth (top) and expected inflation (bottom) in the yield curve formula (14). Taking posterior estimates from* ***tight*** *and* ***joint*** *setting, we form the posterior distribution of* $-D(\tau)/\tau$ *and* $-E(\tau)/\tau$ *and report the median for each prior setting.*

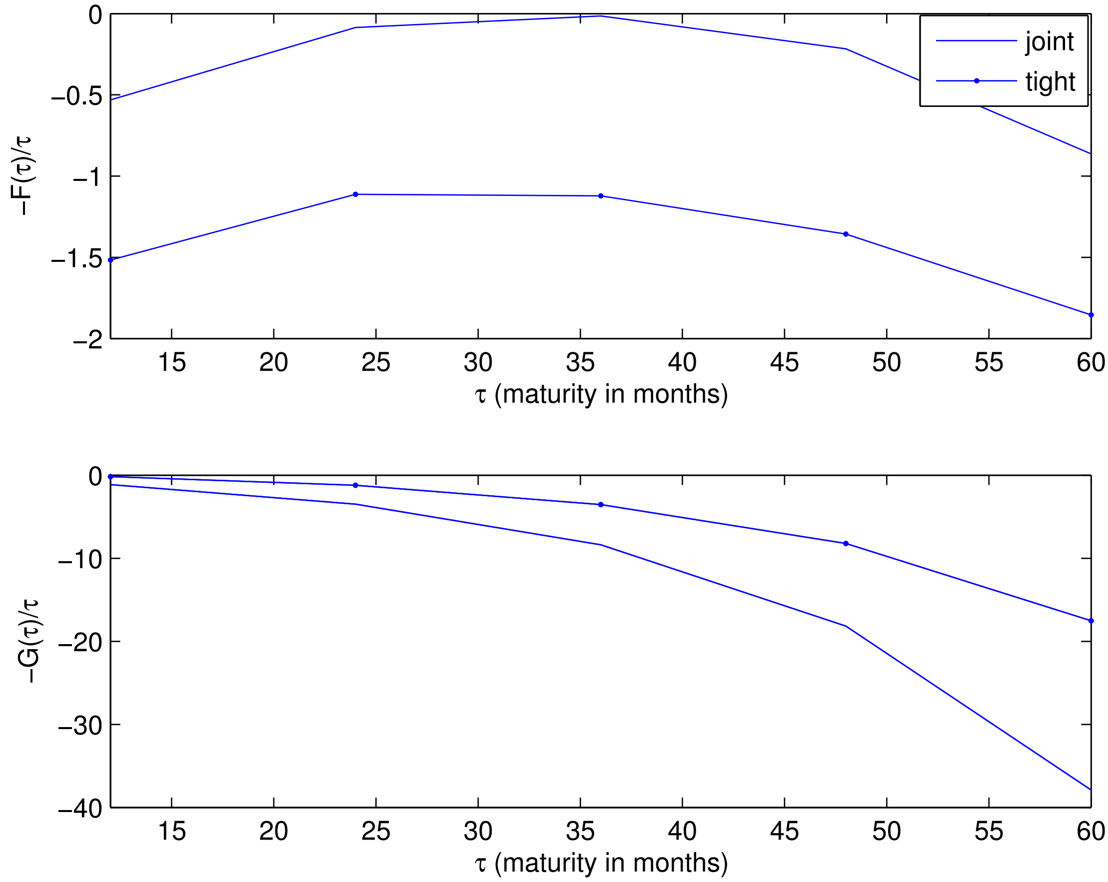

Figure 18: *Loading on the quadratic term of the conditional volatility of expected consumption growth (top) and expected inflation (bottom) in the yield curve formula* (14)*. Taking posterior estimates from* ***tight*** *and* ***joint*** *setting, we form the posterior distribution of* $-F(\tau)/\tau$ *and* $-G(\tau)/\tau$ *and report the median for each prior setting.*

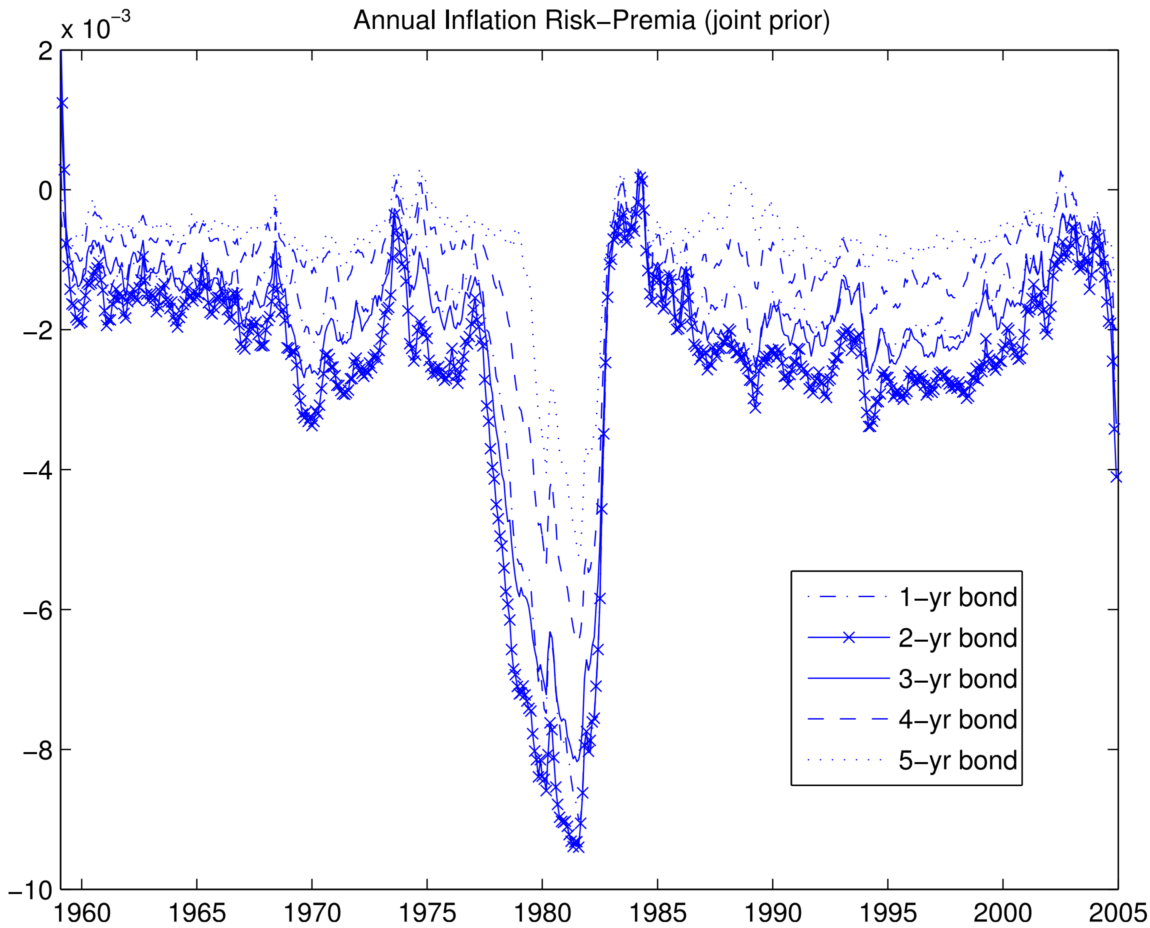

Figure 19: *Inflation-risk premia (IRP) as defined (17) under the* ***joint*** *setting. Taking the posterior distribution of parameters and latent variables under the* ***joint*** *setting, we construct the inflation risk-premia for different maturities and report the posterior median of IRP for each bond.*

5 Appendix

5.1 Appendix A

Proof of Proposition 2.1
The utility function of the investor is

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$$

where $\gamma > 1$. The nominal pricing kernel is defined as

$$M_t^n = e^{-\phi t}\frac{u'(c_t)}{q_t} = e^{-\phi t}\frac{c_t^{-\gamma}}{q_t}$$

Applying Ito's Lemma to M_t^n and using (1)-(2), I get

$$\begin{aligned}\frac{dM_t^n}{M_t^n} &= [-\gamma\mu_t^c - \mu_t^q - \phi + \frac{\gamma(1+\gamma)\sigma_c^2}{2} + \sigma_q^2 + \gamma\rho_{cq}\sigma_c\sigma_q]dt \\ &\quad - \gamma\sigma_c\, dW_c - \sigma_q\, dW_q\end{aligned}$$

Applying $E_t[\frac{dM_t^n}{M_t^n}] = -r_t^n\, dt$ proves the result.

The real interest rate can be obtained by applying the same technique to the real pricing kernel $M_t = e^{-\phi t}c_t^{-\gamma}$. In differential form, the real pricing kernel takes the form

$$\frac{dM}{M} = -[\gamma\mu_t^c + \phi - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2]dt - \gamma\sigma_c\, dW_c$$

with the real rate $r_t = \gamma\mu_t^c + \phi - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2$.
Proof of Proposition 2.2
Given the nominal short rate process in (10), the price of a $\tau = T - t$ period bond, $P^n(\mu_t^c, \mu_t^q, \vartheta_t^c, \vartheta_t^q, \Theta; \tau)$, where Θ is the entire parameter set, is

$$E_t\left[\frac{dP^n}{P^n}\right] = r_t^n dt - E_t\left[\frac{dM^n}{M^n}\frac{dP^n}{P^n}\right]$$

Because of the differential form of the nominal rate process in (10), P^n must be a function of $P^n(\mu_t^c, \mu_t^q, \vartheta_t^c, \vartheta_t^q, \Theta; \tau)$ with the initial condition that $P^n(\mu_t^c, \mu_t^q, \vartheta_t^c, \vartheta_t^q, \Theta; 0) = 1$. An application of Ito's Lemma leads to

$$P^n_{\mu^c}(\alpha^c_0 - \alpha^c_1\mu^c_t - \alpha^c_2\mu^q_t) + P^n_{\mu^q}(\alpha^q_0 - \alpha^q_1\mu^c_t - \alpha^q_2\mu^q_t) + P^n_{\vartheta^c}(\beta^c_0 - \beta^c_1\vartheta^c_t) + P^n_{\vartheta^q}(\beta^q_0 - \beta^q_1\vartheta^q_t) - P^n_\tau$$
$$\frac{1}{2}[P^n_{\mu^c\mu^c}\vartheta^{c^2}_t + P^n_{\mu^q\mu^q}\vartheta^{q^2}(t) + P^n_{\vartheta^c\vartheta^c}\epsilon^2_c + P^n_{\vartheta^q\vartheta^q}\epsilon^2_q] = r^nP^n + \rho_1\gamma\sigma_c P^n_{\mu^c}\vartheta^c_t + \rho_2\sigma_q P^n_{\mu^q}\vartheta^q_t$$

the solution of which is given by (13). After substituting (13) into the above PDE, I can verify the quadratic form and obtain the coefficients which solve a system of ordinary differential equations given by

$$\begin{aligned}
A'(\tau) &= \alpha^c_0 B(\tau) + \alpha^q_0 C(\tau) + \beta^c_0 D(\tau) + \beta^q_0 E(\tau) + \epsilon^2_c F(\tau) + \\
&\quad \frac{1}{2}\epsilon^2_c D(\tau)^2 + \epsilon^2_q G(\tau) + \frac{1}{2}\epsilon^2_q E(\tau)^2 - k \\
B'(\tau) &= -\alpha^c_1 B(\tau) - \alpha^q_1 C(\tau) - \gamma \\
C'(\tau) &= -\alpha^c_2 B(\tau) - \alpha^q_2 C(\tau) - 1 \\
D'(\tau) &= 2\beta^c_0 F(\tau) - \beta^c_1 D(\tau) + 2D(\tau)F(\tau)\epsilon^2_c - \gamma\sigma_c\rho_1 B(\tau) \\
E'(\tau) &= 2\beta^q_0 G(\tau) - \beta^q_1 E(\tau) + 2E(\tau)G(\tau)\epsilon^2_q - \sigma_q\rho_2 C(\tau) \\
F'(\tau) &= -2\beta^c_1 F(\tau) + \frac{1}{2}B(\tau)^2 + 2F(\tau)^2\epsilon^2_c \\
G'(\tau) &= -2\beta^q_1 G(\tau) + \frac{1}{2}C(\tau)^2 + 2G(\tau)^2\epsilon^2_q
\end{aligned}$$

with initial condition $A(0) = B(0) = C(0) = D(0) = E(0) = F(0) = G(0) = 0$. This is a non-linear system of ODEs and is solved in groups using an Euler approximation. First (B, C, F, G) are solved jointly. Then these values are taken to solve (D, E) and finally A.

Real bonds, $P^r(\tau)$ can be priced very similarly using the real interest rate r_t and the real pricing kernel. It will satisfy another set of ODE's, $\hat{A} \cdots \hat{G}$, with boundary condition $\hat{A} = \cdots = \hat{G} = 0$.

5.2 Appendix B

This section details the MCMC methodology that is used in the paper. The entire estimation task can be broken down into filtering the state-variables - μ and v, and parameter estimation Θ. Θ can be broken down into smaller blocks and estimated one at a time. Estimating them separately and one conditional on the other ultimately gives the joint posterior density - $p(\mu, v, \Theta|\text{data})$.

First we look at filtering the growth rates - μ conditional on $\{v, \Theta|\text{data}\}$. The filtering problem looks like

$$\begin{aligned} y_{t+1} &= m_0 + m_1 v_t + m_2 v_t^2 + m_3 \mu_t + W_{t+1} & & W_{t+1} \sim N_3(0, W) \\ \mu_{t+1} &= e + d\mu_t + U_{t+1} & U_{t+1} \sim N_2(0, V_t) & \end{aligned}$$

This is a standard linear filtring problem that can be easily solved by the Forward-Filtering Backward-Sampling (FFBS) algorithm developed in Carter and Kohn (1994) and Fruhwirth-Schnatter (1994), except with one twist. The error terms W_{t+1} and U_{t+1} are correlated with correlation Σ_t as is defined in Section 3.1. Thus, when we perform the FFBS we need to be mindful of this correlation structure which is not there in the original FFBS. We orthogonalize the two equations using $U_{t+1}|W_{t+1}$ which is still normally distributed and use this conditional distribution in the bottom equation and proceed with standard FFBS.

Having obtained the time-series of the full growth rates μ_t, now we move on to estimating the volatility conditional volatility v_t. Notice the state-space is highly non-linear in the volatility states v_t.

$$\begin{aligned} y_{t+1} &= m_0 + m_1 v_t + m_2 v_t^2 + m_3 \mu_t + W_{t+1} & & W_{t+1} \sim N_3(0, W) \\ \mu_{t+1} &= e + d\mu_t + U_{t+1} & U_{t+1} \sim N_2(0, V_t) & \\ v_{t+1} &= f + g v_t + E_{t+1} & E_{t+1} \sim N_2(0, E) & \end{aligned}$$

It is quadratic in the observation equation to reflect that the yield curve is quadratic in the volatility states and highly nonlinear in the first state-equation. Appendix C generalizes the arguments of FFBS to filter *and smooth* the volatility states.

Now we have to estimate the parameter set in Θ. It comprises of regression parameters - (e, d) and (f, g), volatility parameters - W and E, correlation parameters ρ_1 and ρ_2, and utility parameters γ and ρ. Consider the full posterior of (e, d).

$$p(e, d|\Theta_{-(e,d)}, \mu, v, \text{data}) = p(e, d)p(\mu|v, (e, d), \Theta_{-(e,d)})p(\text{data}|, (e, d)\Theta_{-(e,d)}, \mu, v)$$

When the yields are included in the data, there are no closed-form posterior densities available for (e, d). We follow a random-walk Metropolis algorithm which proceeds in the following sense.

- 1. Start with a given (e^0, d^0).
- 2. Propose a new candidate $(e^1, d^1) \sim N((e^0, d^0), S)$ where S is chosen to be very "small" so that the chain makes very small progress.
- Accept this new point as the new draw from the posterior distribution with probability $a = \min\left[1, \frac{p(e^1,d^1|\Theta_{-(e^1,d^1)},\mu,v,\text{data})}{p(e^0,d^0|\Theta_{-(e^0,d^0)},\mu,v,\text{data})}\right]$

Without the yields, a closed form conjugate prior is multivariate normal which produces posterior distribution of the regression coefficients to be multivariate normal. The exact specification is given in Johannes and Polson (2005).

Next, we move on to the volatility of volatility parameters. As usual, in the presence of bond prices, a closed form solution is not available. In fact, no good proposals are also available. Thus, we had to resort to a grid, popularly known as Griddy Gibbs algorithm. We fix a grid for a volatility parameter $\epsilon_c - \{\epsilon_c^1, \cdots, \epsilon_c^n\}$ with a flat prior on the grid. At each point on the grid, we evaluate the posterior distribution,

$$p(\epsilon_c^i|\Theta_{-\epsilon_c^i}, \mu, v, \text{data}) = p(\epsilon_c^i)p(v|\epsilon_c^i, \Theta_{-\epsilon_c^i})p(\text{data}|\epsilon_c^i, \Theta_{-\epsilon_c^i}, \mu, v)$$

Then draw ϵ_c from this grid of densities $\{p(\epsilon_c^1), \cdots, p(\epsilon_c^n)\}$. As usual, without the yield data a conjugate prior is inverted-chi square which produces a closed form posterior distribution for ϵ_c to be inverted chi-square. The exact specification is given in Johannes and Polson (2005). The same method applies to ϵ_q. Other parameters, like the correlations ρ_1, ρ_2 and the utility parameters γ and ϕ have no known posteriors. Whereas γ and ϕ can only be inferred from the bond data, ρ_1 and ρ_2 can be obtained from both the yield curve and the macroeconomic growth rates. They are sampled via the same Griddy Gibbs algorithm.

The next step is to draw the lower partition of the W-matrix. Let $W' = \left[\sigma_c^2 \;\; \rho\sigma_c\sigma_q; \rho\sigma_c\sigma_q \;\; \sigma_q^2\right]$. The posterior distribution is the same as before,

$$p(W'|\Theta_{-W'}, \mu_t, v_t, \text{data}) = p(W')p(\text{data}|W', \Theta_{-W'}, \mu_t, v_t, \text{data})$$

Without the bond data a conjugate prior for W' is inverted Wishart. With the bond data, a choice for the proposal density is given by the macroeconomic time series

$$W'^{\text{next}} \sim p(W'^{\text{previous}})p(y^c, y^q|W'^{\text{previous}}, \Theta_{-W'^{\text{previous}}})$$

which is a draw from the closed-form posterior and the acceptance probability given by $a = \min\left[1, \frac{p(W'^{\text{next}}|\text{yield}\cdots)/q(W'^{\text{next}}|W'^{\text{previous}})}{p(W'^{\text{previous}}|\text{yield},\ldots\cdots)/q(W'^{\text{previous}}|W'^{\text{next}})}\right]$, where q is the proposal density above.

5.3 Appendix C

Generalized FFBS algorithm to filter volatility

The generalized version of the non-linear Kalman filter that applies to any level of non-linearity in the observation or transition density is presented in Hore, Lopes and McCulloch (2009). Here, we provide a specific example of their algorithm adopted to fit the non-linear filtering exercise involved in the estimation of the macroeconomic conditional volatility v_t^c (v_t^q) from the quadratic yield curve.

A univariate version of the filtration of conditional volatility is described here.

$$
\begin{aligned}
X^1_{t+1} &= c_1 v_t + c_2 v_t^2 + kZ_0 && (21)\\
X^2_{t+1} &= v_t Z_1 && (22)\\
v_{t+1} &= \alpha + \beta v_t + cZ_2 && (23)
\end{aligned}
$$

where Z_0, Z_1 and Z_2 are $N(0,1)$ and independent of each other. The first equation (21) is similar to the yield curve where the conditional volatility enters as a quadratic with loading c_1 and c_2 that are known given all the other parameters. The second equation (22) reflects the state-equation for expected consumption growth or expected inflation where the conditional volatility enters in the error term.

Let's say the volatilities v_t can only take discrete values $Y = \{y_1, \ldots, y_n\}$. Depending on the problem, one has to pick a grid that is appropriate for the problem with positive support and reaches a maximum within the grid.

Assume that the prior probabilities on v_0, is given, which means that for state 0, $p(Y) = \{p(y_1), \ldots, p(y_n)\}$ is known. The goal is to create the posterior probability of state 1, $p(v_1|X^1_1 = a, X^2_1 = b)$.

Notice that the conditional distribution of v_1 is given by the state equation (23)

$$
p(v_1 = y_j | v_0 = y_k) \sim N(\alpha + \beta y_k, c\sqrt{y_k}) \tag{24}
$$

where $y_j, y_k \in Y$. This represents the transition probabilities for the volatilities going from state y_k to state y_j.

The joint probability of $p(v_1, v_0)$ can be obtained from Bayes' rule

$$
p(v_1 = y_j, v_0 = y_k) = p(v_1 = y_j | v_0 = y_k) p(v_0 = y_k) \tag{25}
$$

The first term on the right is the transitional probability from state y_k to state y_j that one can obtain from (23). The second equation is the prior probability distribution which is given.

In order to get the posterior distribution of $p(v_1|X^1_1 = a, X^2_1 = b)$, first compute the full joint distribution.

$$
\begin{aligned}
p(X^1_1 = a, X^2_1 = b, v_1 = y_j, v_0 = y_k) &= p(X^1_1 = a, X^2_1 = b | v_1 = y_j, v_0 = y_k)\\
&\quad\; p(v_1 = y_j, v_0 = y_k)\\
&= p(X^1_1 = a | v_1 = y_j) p(X^2_1 = b | v_1 = y_j)\\
&\quad\; p(v_1 = y_j | v_0 = y_k) p(v_0 = y_k) && (26)
\end{aligned}
$$

The first two terms on the right can be computed from (21)-(22). The last two terms represent the joint probabilities discussed above.

Now, following Bayes' rule again the joint distribution of v_1, v_0 conditional on $(X_1^1 = a, X_1^2 = b)$ can be written as

$$p(v_1 = y_j, v_0 = y_k | X_1^1 = a, X_1^2 = b) = \frac{p(X_1^1 = a, X_1^2 = b, v_1 = y_j, v_0 = y_k)}{p(X_1^1 = a, X_1^2 = b)} \tag{27}$$

The numerator is determined by (26). The denominator is a constant conditional on $(X_1^1 = a, X_1^2 = b)$. Hence, the above can be written as

$$p(v_1 = y_j, v_0 = y_k | X_1^1 = a, X_1^2 = b) \propto p(X_1^1 = a, X_1^2 = b, v_1 = y_j, v_0 = y_k) \tag{28}$$

Finally, integrate out v_0 in order to get the posterior distribution of $v_1 | X_1^1 = a, X_1^2 = b$.

$$p(v_1 = y_j | X_1^1 = a, X_1^2 = b) = \sum_{v_0 \in Y} p(X_1^1 = a, X_1^2 = b, v_1 = y_j, v_0) \tag{29}$$

Let $D_t = \{(X_1^1, X_1^2), \ldots, (X_t^1, X_t^2)\}$. Now, use $p(v_1 | D_1)$ as the prior information to get $p(v_2 | D_2)$ and so on.

Filter forward in this way until the last observation (X_T^1, X_T^2) is reached and sample v_T right away from $p(v_T | D_T)$. The backward filtering requires that the state one period ahead is known. So, once v_T, is filtered, draw v_{T-1} all the way back to v_1.

$$p(v_t = y_j | v_{t+1}, D_t) = \frac{p(v_{t+1} | v_t = y_j) p(v_t = y_j | D_t)}{\sum_{v_t \in Y} p(v_{t+1} | v_t) p(v_t | D_t)} \tag{30}$$

The first term on the numerator (denominator) is the transition probability in (23) while the second term is the obtained from the forward filtering (29).