Filtering Macroeconomic Variables from A Non-Linear Term Structure Model

Satadru Hore

November 25, 2005

1Graduate School of Business, University of Chicago, Chicago, IL 60637. E-mail: pshore@gsbphd.uchicago.edu. I would like to thank my committee: Alan Bester, George Constantinides, Robert McCulloch (Chairman) and Pietro Veronesi. I would also like to thank John Cochrane, Nick Polson, Hedibert Lopes, Hui Chen and seminar participants in asset pricing brownbag in Chicago. For disclosure purposes, all computation was carried out in Linux platform using GNU Scientific Library. All errors are mine.
Abstract

An integrated utility-based model is presented for both real and nominal bond prices, where CPI and real consumption follow a jointly lognormal process with unobserved growth rates. The volatilities of the growth rates are also stochastic which give rise to time-varying risk-premia. The model is fitted to the time-series of fundamentals in the economy and the yield curve under different prior settings that consider a varying mix of bi-directional information between no-arbitrage yields and underlying macroeconomic dynamics. The prior specification that considers a moderate mix of no-arbitrage with underlying fundamentals does a superior job in predicting in-sample future yields and also matches the risk-free rate out of sample. The prior specification that specifies a heavier emphasis on no-arbitrage bond prices do worse in predicting future yields and fitting macroeconomic dynamics, but can generate inflation risk-premia observed in the literature.
1 Introduction

Traditional dynamic term structure models in finance model the short rate process as a linear combination of a small number of factors that are latent (Dai and Singleton (2000), Bauduzzi et al. (1996), Chen and Scott (1993)). However, the resulting arbitrage-free model of bond prices have poor forecasting characteristics (Duffee (2002), Dai and Singleton (2003)). This paper shows that one-period ahead forecasting of yields is best achieved with a blend of macroeconomic variables as well as arbitrage-free pricing. This is accomplished within an utility based no-arbitrage model of bond prices with exogenously given consumption and CPI growth, along with the dynamics of the volatilities of expected consumption growth and expected inflation. The Bayesian setting allows for consideration of a wide range of models that allow for varying weights given to either the macroeconomic setting, or to the yield curve through the choice of prior information. As such, it can generate a reduced form term-structure model as well as produce models that can try to fit both the macro series and the yield curve.

The model starts with an exogenous specification of growth rates of consumption and CPI, where the growth rates are unobservable and follow a linear system with stochastic volatility, whose processes are also exogenously specified. The stochastic nature of the volatilities give rise to bond risk-premia that are time-varying - a key feature of the term-structure. A contribution of this paper is that it endogenizes time-varying bond risk-premia given an endogenous pricing kernel within a macroeconomic setting. A key proposition that can be examined directly from this setting is whether a nominal term-structure model based on economic fundamentals carry information about latent factors governing the fundamentals that are different from a solo analysis of the fundamentals. In fact, the recent macro term-structure literature has started exploring this issue.

In recent years, macroeconomists have proposed several models of the yield curve - with or without the no-arbitrage restriction. Diebold and Li (2005) have shown that a latent three factor extension (labelled “level”, “slope” and “curvature”) of the Nelson-Siegel paradigm produces better forecasts of the yield curve than standard benchmark models. Researchers have also been incorporating macroeconomic observables in addition to latent factors to study the yield curve. Diebold, Rudebusch and Aruoba (2005) incorporate the Nelson-Siegel representation within standard macroeconomic VAR, while Ang and Piazzesi (2003) incorporate it within a no-arbitrage setting. A bi-directional interaction among parameters (i.e. parameters jointly governing macro processes and yields) is also considered in Ang, Dong and Piazzesi (2004) and Diebold, Rudebusch and Aruoba (2005).

This paper employs a Bayesian methodology that takes the bi-directional approach another step by considering different prior specifications. On one extreme, a choice of prior can fit yields discarding the macroeconomic series and have high value of risk-aversion, and on the other extreme, can fit the macro series well by foresaking the yields and generate very low risk-aversion parameter. The latter is consistent with an empirical calibration of
the risk-free rate in standard Euler equation setting. However, both of these specifications generate in-sample forecasts of yields that are inferior to the prior setting where there is a mixture of bond yields and macroeconomic dynamics. What is also revealing in this setting is that a solo look at the yield curve will provide volatilities in expected consumption growth and expected inflation that are simply not observed in the underlying macro dynamics. But, a better mix of fundamentals and no-arbitrage model of yields produce intermediate levels of volatilities between the “quasi”-reduced form specification and macro factors.

The processes for the latent factors governing the drift of consumption and inflation are borrowed from the statistical term-structure literature. There is a feedback relationship between the expected growth rates, such that the change in expected consumption growth rate depends on expected inflation, and vice versa. This feedback relationship is found in the maximal model of Dai and Singleton(2003), where the effective purpose was to introduce feedback relationships between the central tendency and volatility factors of the short rate. In this case, its effect is that the expected change in the real rate is affected by expected inflation. This is a violation of the Fisher hypothesis, which has been documented by Pennacchi(1991) and Goto and Taurus(2005), where they have found a positive correlation between expected real rates and expected inflation. Once again, an unidirectional approach of analyzing the fundamentals only does not generate that effect as well as a joint analysis of economic fundamentals and macro variables.

The latent growth rates of consumption and CPI along with their volatilities form the vector of state variables which prices a nominal or real bond with any maturity. What links the state variables to the bond prices is a set of factor loadings which are derived from the utility based no-arbitrage approach plus a Jensen’s inequality term. Essentially, it is the same as a Nelson-Siegel setting where the factors have conditional stochastic volatility and the factor loadings for bond prices are not specified directly but derived endogenously under the factor dynamics and no-arbitrage.

An innovation in this work is the introduction of stochastic volatility into the drift rates, which yields bond prices that fall into the Quadratic-Gaussian family. A straight-forward inversion of the yield curve to obtain volatility estimates is fairly cumbersome, and a non-linear filtering methodology is implemented to invert the problem and obtain the volatility parameters. This method is also implemented keeping in mind the bi-directional approach in the literature. Volatilities must be obtained keeping in mind both the yield curve and the macro variables. Once again, different prior specification can yield different time-series of volatilities which can be compared against the actual volatility of the fundamentals. This method is also new and is jointly developed with Robert McCulloch.

Literature Review
The work that is closest to this paper is by Veronesi and Yared(2000) and Pennacchi(1991). In Veronesi and Yared, the drift rates of real consumption and CPI undergo regime shifts, while in this model the drifts follow an Ornstein-Uhlenbeck process with stochastic volatility. Pennacchi’s model is the closest to this one - it has interaction between real interest
rates and inflation, but the risk premia is constant. Also, he assumes that investors have log-utility which precludes the estimation of risk preference parameter and also restricts the size of the risk premia and the real interest rate. Moreover, he does not include consumption within his state-space and restricts himself to looking at the joint distribution of interest rates and inflation. Other closely related works on real and nominal interest rates are by Evans (1998, 2003), Buraschi and Jiltsov (1999) and Sun (1992). Evans (2003) develops a Markov switching model, but restricts the term structure within an affine class model, thus restricting the preference of the investor. Both Evans and Pennacchi estimate their state variables from a state-space system, which requires the assumption of a pricing error, i.e. the model prices the yield curve up to an error. Instead of using real bonds, as in Evans (1998), the model here considers the joint processes of CPI, real consumption and nominal rates and can back out the real rate.

Recent works in incorporating macro variables in no-arbitrage models of the yield curve include Piazzesi (2005), where the observable factor is the Federal Reserve’s rate for overnight lending. Ang, Piazzesi and Wei (2005) jointly model the dynamics of yields and GDP growth. Diebold, Rudebusch and Aruoba (2005) interpret the correlations Nelson-Siegel yield factors and macroeconomic variables. Diebold, Piazzesi and Rudebusch (2005) gives us an affine interpretation of the Nelson-Siegel representation, where the latent variables can be related to macroeconomic factors.

The paper is organized as follows: The next section presents the model. The following section describes the state-space system and describes the estimation method. That is followed by a section on discussion of the empirical findings and finally concludes.

2 The Model

Let real consumption $c_t$ and price of the consumption good (CPI) $q_t$ jointly follow a geometric Brownian motion process

$$\frac{dc_t}{c_t} = \mu_c dt + \sigma_c dW_c(t)$$ (1)

$$\frac{dq_t}{q_t} = \mu_q dt + \sigma_q dW_q(t)$$ (2)

where $\text{corr}(dW_c(t), dW_q(t)) = \rho_{cq} dt$. Both series follow a geometric processes with growth rates that are changing over time but with constant volatility and correlation. These growth rates are not observable to the investor and the exact nature of them will be considered later. Veronesi (2000) assumed these growth rates take on discrete values from some set following a jointly Markov structure. In the model, macroeconomic information is captured in the percentage growth in consumption and CPI, which shall be referred to as fundamentals.
Let the investor have an utility function of the form \( u(c_t) = \frac{1}{1-\gamma} \). Furthermore, let the nominal pricing kernel be defined by \( M^n_t = e^{-\phi_t \frac{u'(c_t)}{c_t}} \). It is the usual pricing kernel in continuous time denominated by the price of the consumption good in order to make it nominal. Given an utility function, the pricing kernel, the consumption \( (1) \) and price processes \( (2) \), the nominal short rate process for the investor follows:

**Proposition 2.1** The nominal short rate process is

\[
    r^n_t = \gamma \mu_c^t + \mu_q^t + k
\]

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

\[
    r_t = \gamma \mu_c^t + \phi - \frac{1}{2} \gamma (1 + \gamma) \sigma_c^2
\]

**Proof:** In Appendix.

The nominal interest rate in Proposition (2.1) is linear in the latent drifts of consumption and CPI that are unobservable to the investor. However, the processes of the drift rates are known to her. The latent growth rates can be either positive, negative or zero. A Gaussian model of the drift rates make sense in this situation but it will have a very similar drawback as it’s Vasicek(1977) counterpart, namely it will generate negative interest rates. While a negative real rate is very likely, ideally the nominal rate should be restricted to be positive. However, this leads to significant difficulties in deriving equilibrium bond prices and this restriction is not imposed. In the three-factor term-structure models of Dai and Singleton(2000), the family of affine diffusions that guarantee strict positivity are multi-factor square-root models, among which negative correlations are ruled out, thereby restricting the state dependence of the market prices of risk. Admissibility of these models (i.e. they generate positive interest rates) is enforced by restricting the parameter space of the models. However, the restrictions that are necessary are not binding and the estimated parameters of the model generate positive nominal rates almost surely. The nominal rate process can also be interpreted in the traditional Nelson-Siegel two-factor setting where the factors are expected growth rates of consumption and inflation, plus a constant term which reflects precautionary savings motive of the investor.

The set of latent growth rates follow an interdependent process, such that the expected growth rates of both processes affect each other. Its implication is that the expected change in the real rate depends on expected inflation. In other words, it reflects the failure of the Fisher hypothesis as pointed out by Evans(1998) and Pennacchi(1991). Furthermore, the conditional volatility of these drift rates are stochastic.
The structure of the drifts is the following:

\begin{align*}
\mathrm{d}\mu_t^c &= (\alpha_0^c - \alpha_1^c \mu_t^c - \alpha_2^c \mu_t^q) \, \mathrm{d}t + \theta_t^c \, \mathrm{d}B_c(t) \\
\mathrm{d}\mu_t^q &= (\alpha_0^q - \alpha_1^q \mu_t^c - \alpha_2^q \mu_t^q) \, \mathrm{d}t + \theta_t^q \, \mathrm{d}B_q(t)
\end{align*}

where the Brownian motions \(\mathrm{d}B_c\) and \(\mathrm{d}B_q\) are independent. The terms \(\alpha_2^c\) and \(\alpha_2^q\) are the feedback terms which impact each other’s future growth rates. Rewriting the real rate from (2.1) in differential form,

\[ \mathrm{d}r_t = (a(t) - \alpha_1^c r_t^c) \, \mathrm{d}t + \gamma \theta_t^c \, \mathrm{d}B_c(t) \]

where \(a(t) = \gamma \alpha_0^c + \alpha_1^c (\phi - \frac{1}{2} \gamma (1 + \gamma) \sigma_c^2) - \gamma \alpha_2^q \mu_t^q\). Here the real rate follows an autoregressive process with a time-varying central tendency which is determined by the current rate of inflation. If \(\alpha_2^c\) is positive, the higher the rate of inflation, the lower the long term central tendency of the real rate process. Despite many assumptions about independence between real rate and expected inflation, recent evidence shows positive relationship between the two implying \(\alpha_2^c < 0\). However, this result will be opposite under expected deflation. Also, a higher risk-aversion parameter lowers the central tendency because of higher precautionary savings if higher risk-aversion can overcome the positive influence of expected inflation. This setup captures the failure of the Fisher hypothesis since the conditional mean of future real rate depends on expected inflation and also directly maps into the essentially affine Duffee(2002) specification of market prices of risk in the term structure literature. If future consumption and present inflation are negatively correlated, then this predicts the future real rate to be lower and higher if today’s state suggests deflation. A similar differential form can also be written for the nominal rate.

\[ \mathrm{d}r_t^n = (a^n(t) - \alpha_1^c r_t^n) \, \mathrm{d}t + \gamma \theta_t^c \, \mathrm{d}B_c(t) + \theta_t^q \, \mathrm{d}B_q(t) \]

where

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

This is similar to the term structure models of Chen(1996) and Balduzzo et al.(1998,1996). In this setting, the central tendency is determined by expected inflation and the real rate, and the volatility of the short rate is determined by both the volatilities of the drift rates of consumption and CPI. The effect of real rate on the long run mean of the nominal rate is controlled by the parameter \(\alpha_1^q\). Whereas the statistical term structure literature starts with (8), the nominal rate derived here is based on the preference of the investor along with some assumption of the fundamentals of the economy. This short rate is also similar to Collin-Dufresne et al.(2004) who found a four-factor(one of which determines the conditional volatility of all four state variables) unspanned stochastic volatility model to
generate both realistic volatility and a good cross-section fit under a form of market-price of risk that is a generalized form version of Duffee’s (2002) essentially affine specification. This model falls under the category of a four-factor unspanned stochastic volatility model where two of the factors are controlling the volatility of the nominal rate, however the model here is clearly not maximal.

Technically, we do not need to restrict the sign of \( \{ \vartheta^c_t, \vartheta^q_t \} \), because even with a negative value of \( \vartheta^c_t \), the quadratic variation of \( \mu^c_t \) is positive, i.e. \( E[(d\mu^c_t)^2] = (\vartheta^c_t)^2 \geq 0 \). Hence, modelling them as arithmetic processes is mathematically sufficient. However, in order to accommodate the right-sign of the risk-premia, it maybe necessary to restrict the sign of the volatility processes, which is done by the choice of priors. As suggested by Stein and Stein (1991), under a wide set of parameter values the probability of \( \vartheta^c_t \) ever becoming less than or equal to zero is very small.

The process the \( \vartheta^i \)'s follow is

\[
d\vartheta^i_t = (\beta^i_0 - \beta^i_1 \vartheta^i_t) \, dt + \epsilon^i_t \, dZ^i_t
\]

for \( (i = c, q) \) where \( dZ^i \)'s are uncorrelated with each other and with the other Brownian terms. The admissibility of the model of bond prices resulting from this setting is contingent on positive volatilities and the parameters of the system giving rise to positive nominal rates.

**Bond prices:** Having specified the fundamentals of the economy, the equilibrium no-arbitrage price of a bond \( P(r^n_t, \tau) \) at time \( t \) with maturity \( \tau \) is given by the following relationship

\[
E_t(dP) = r^n_t P + E_t \left( \frac{dM^n}{M^n} \, dP \right)
\]

The first term on the right is the usual nominal return of a bond for an investor who doesn’t need to be compensated for holding the bond. The second term reflects the risk-premia term whose counterpart in the term-structure literature is the market price of risk that is generally exogenously specified but here is identified by preferences and the processes of consumption and CPI. The correlations needed in order to generate time-varying risk-premia are \( \rho_1 = corr(dW^c, dB^c) \) and \( \rho_2 = corr(dW^q, dB^q) \). Notice that a correlation between the volatilities and the consumption/inflation processes can also be accommodated within an affine framework given arithmetic processes like (9), but its effect would only increase the risk-premia by a constant. Correlation between the volatilities will increase the complexity of the problem much more by introducing a cross-product term.

**Proposition 2.2** 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 (5)-(6) and volatility given by (9) is

\[
P^n_t = \exp\left(A(\tau, \Theta) + B(\tau, \Theta)\mu^c_t + C(\tau, \Theta)\mu^q_t + D(\tau, \Theta)\vartheta^c_t + E(\tau, \Theta)\vartheta^q_t + F(\tau, \Theta)\vartheta^c_t^2 + G(\tau, \Theta)\vartheta^q_t^2\right)
\]
where $\Theta$ is the entire parameter space and $A \cdots G$ constitute a system of ordinary differential equations with initial condition $A(0) = \cdots = G(0) = 0$.

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

**Proof:** In Appendix.

The difference between real and nominal bonds in this setting arises from the difference in risk premia and in the short rate. The difference between the two short rates is discussed in (2.1). The real pricing kernel has only one source of risk - from the volatility of expected consumption growth, and therefore only compensates the bond holder from that source of risk. There is no inflationary risk as the real bond delivers the monetary equivalent of one unit of consumption good.

The bond prices that come out of this setting is a member of the Quadratic-Gaussian family, a special case of which is the SAINTS model proposed by Constantinides(1992). The stochastic volatility gives rise to terms that are both linear and quadratic in those volatilities. During the empirical analysis, an error term will be added to the bond pricing equation so that the unobservable states can be filtered out. The use of a pricing error has been used widely in Evans(2003), Campbell and Viceira(2001), Polson and Johannes(2003) and others. The intuition behind the error term is that clean prices may not be observed, but the deeper econometric reason is discussed later.

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

\[
E_t \left[ \frac{dP}{P} \right] = (r_t^n - \rho_1 \gamma \sigma_c B(\tau) \varphi_t^c - \rho_2 \sigma_q C(\tau) \varphi_t^q) \, dt \\
E_t \left[ \frac{d\hat{P}}{\hat{P}} \right] = (r_t - \rho_1 \gamma \sigma_c \hat{B}(\tau) \varphi_t^c) \, dt
\]

The ODEs are not solved in closed form because the system is highly non-homogeneous with time-varying coefficients. A fully closed form solution will not yield any new intuition. However, the signs of the solutions should be of immense importance. Since $B$ and $C$ are the loadings on the drift rates, it is safe to conjecture that they should be negative. While a higher rate of inflation must reduce bond prices (i.e. raise interest rates), a higher growth rate of consumption also increases the nominal rate as long as $\gamma > 0$.

The nominal bond return exhibits stochastic volatility because of the stochastic volatility in the growth rates, and its conditional variance is quadratic in the volatilities of consumption and CPI growth rates. The conditional variance of the nominal bond return in
this setting is

\[
V_i \left[ \frac{dP^n}{P^n} \right] = [D(\tau)^2 + E(\tau)^2] + 4F(\tau)D(\tau)v_i + 4G(\tau)E(\tau)v_i^2 + \\
[B(\tau)^2 + 4F(\tau)^2)v_i^2 + [C(\tau)^2 + 4G(\tau)^2)v_i^2]
\] (14)

**Inflation Risk Premia:** The inflation risk premia of a \( \tau \) period bond can be computed as the difference between the excess returns of a nominal and a real bond. The presence of inflation risk-premia would suggest the failure of risk-neutrality and expectations hypothesis. In this case, the inflation risk-premia for a bond with maturity \( \tau \) is

\[
IRP(\tau) = \left( E_i \left[ \frac{dP}{P} \right] - r_i^n \right) - \left( E_i \left[ \frac{dP}{P} \right] - r_i \right)
\]

\[
= -[\rho_1 \gamma \sigma_c (B(\tau) - \hat{B}(\tau))\phi_i + \rho_2 \sigma_q C(\tau)\phi_i]
\] (15)

The IRP in (15) is linear and state-dependent in the volatilities of the growth rates. Instantaneously, the difference between the rates is

\[
r_i^n - r_i = \mu_i^q - \sigma_i^2 - \rho c_\gamma \sigma_c \sigma_q
\] (16)

From this relationship, the difference between the nominal and the real rate could be negative under deflation, as long as the correlation between the errors in consumption growth and inflation is not too large.

### 3 Estimation

The estimation of the parameters and the state variables of the model is done using a Markov Chain Monte Carlo (MCMC) method called Metropolis-Hastings (MH) algorithm. This methodology is used since it provides us with a powerful tool to sample from the posterior distribution of the parameter space even when the exact posterior distribution of the parameters is not precisely known. There are several important theoretical underpinnings of the MCMC algorithm.

First, the Hammersley-Clifford theorem, which states that the joint distribution of the parameters can be completely specified by considering the full set of conditional distributions. Second, the convergence of the posterior distribution is based on the ergodic theory for Markov Chains. Posterior draws generated by MH algorithms have special properties which allow for convergence. These posterior draws are time-reversible and have the target
distribution as an invariant distribution. Aperiodicity for MCMC methods is proven by Tierney(1994).

To proceed with the empirical investigation, first the continuous time-model is discretized to one-month interval, such that \( dt = 1 \) month. Let a \( \tau \) period bond yield at time \( t \) be defined by \( y_t^{P(\tau)} = -\log P_t^{P(\tau)} \). Let there be \( n \) such maturities \( \tau_1, \cdots, \tau_n \). Percentage consumption and CPI change at time \( t \) be defined by \( y_c^t = \frac{c_{t+1} - c_t}{c_t} \) and \( y_q^t = \frac{q_{t+1} - q_t}{q_t} \). Let

\[
y_t = \begin{bmatrix}
y_t^{P(\tau_1)} \\
\vdots \\
y_t^{P(\tau_n)} \\
y_t^c \\
y_t^q
\end{bmatrix}, \mu_t = \begin{bmatrix}
\mu_t^c \\
\mu_t^q
\end{bmatrix} \text{ and } v_t = \begin{bmatrix}
v_t^c \\
v_t^q
\end{bmatrix}.
\]

Also, let \( \Theta \) represent the entire parameter space for the model without the state variables. So, \( \Theta \) contains all the regression, volatility, risk-aversion and discount rate parameters. Some of the elements of \( \Theta \) are scalars while others are multidimensional. Let \( \Theta_{-p} \) denote the entire parameter set without the \( p \)'th parameter. The flexibility of this approach is that it allows for a joint estimation of \((\Theta, \mu, v|y)\) from its equilibrium distribution using both the yield curve data and the economic series.

Consider a partition of \( \Theta \) in \( r \) components \( \Theta = (\Theta_1, \cdots, \Theta_r) \), where each component \( \Theta_j \) could be multidimensional. Given a partition, the Hammersley-Clifford theorem implies that the following set of conditional distributions

\[
\Theta_j|\Theta_{-j}, y, \mu, v \quad \forall j = 1, \cdots, r
\]

uniquely determines \( p(\Theta|\mu, v, y) \). The MH algorithm now boils down to choosing a set of initial values for \((\Theta^0, \mu^0, v^0)\), then draw \( \Theta_j^0|\Theta_{-j}^0, \mu^0, v^0, y \) and finally draw \( \mu, v|\Theta, y \). Then keep repeating the steps until the algorithm is drawing from the equilibrium distribution. The last step is discussed first where the state variables are filtered conditional on knowing all the parameters \( \Theta \).

### 3.1 The state-space system

In all there are 4 state variables in this model, but there are \( n + 2 \) observable data series - \( n \) being the number of bonds that are considered in the sample, and per-capita real consumption and CPI. If \( n > 2 \), then the number of observables at each date is greater than the number of factors. One may assume that two of those bonds are priced perfectly, and can invert the system to obtain the state variables. Duffie and Singleton (1997), for example, assume that the 2 and 10-year swap yields are priced perfectly by their two-factor model. However, the parameter and state variable estimates will be heavily biased.
by the points chosen and may do an inadequate job for other yields. In this setting, a measurement error for yields is added to convert the theoretical model into a state-space system and a filtering methodology is used to back out the state variables from the yields and macroeconomic data.

The filtering problem is further split into two so that first $\mu|v,y,\Theta$ is drawn and then $v|\mu,y,\Theta$ is drawn. First, the $\mu$’s. Fix the following notation:

$$m_t = \begin{bmatrix} A(\tau_1) + D(\tau_1)v_1^2 + E(\tau_1)v_1^2 + F(\tau_1)v_1^2 + G(\tau_1)v_1^2 \\ \vdots \\ A(\tau_n) + D(\tau_n)v_n^2 + E(\tau_n)v_n^2 + F(\tau_n)v_n^2 + G(\tau_n)v_n^2 \\ 0 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} B(\tau_1) \\ \vdots \\ B(\tau_n) \\ 1 \\ 0 \end{bmatrix}, \quad e = \begin{bmatrix} \alpha_0^c \\ \alpha_0^q \end{bmatrix},$$

$$V_t = \begin{bmatrix} \vartheta^2 & 0 \\ 0 & \vartheta^2 \end{bmatrix}, \quad W = \begin{bmatrix} \epsilon^2 I_{nX_n} & 0 & 0 \\ 0 & \sigma_c^2 & \rho\sigma_c\sigma_q \\ 0 & \rho\sigma_c\sigma_q & \sigma_q^2 \end{bmatrix}, \quad d = \begin{bmatrix} (1 - \alpha_1^c) & -\alpha_2^c \\ -\alpha_1^q & (1 - \alpha_2^q) \end{bmatrix}$$

$$\Sigma_t = \begin{bmatrix} \rho_1 \vartheta^2 \sigma_c & 0 & 0 \\ 0 & \rho_2 \vartheta^2 \sigma_q & \rho_2 \vartheta^2 \sigma_q \end{bmatrix}.$$

The bond yield, consumption and CPI growth can be written in the familiar state-space system,

$$y_t = m_t + b\mu_t + W_t \quad W_t \sim N_3(0,W)$$

$$\mu_{t+1} = e + d\mu_t + U_t \quad U_t \sim N_2(0,V_t)$$

This notation is extremely flexible. The states ($\mu_t$) can be filtered with the bonds by considering the full-system, and without the bonds by simply eliminating the top element of $y_t$ and its corresponding elements in $(m_t, b, W)$. The role that prior plays in this analysis can be observed through the variance-covariance matrix $W$. $W$ is a block diagonal matrix, with the upper left block representing the exogenous error term imposed upon the system for the purpose of filtering. The lower right block is the variance-covariance matrix of the consumption/CPI-growth system. If a strong prior is imposed on the yield-error $\epsilon$, then the state variables filtered from the system will fit the bonds more closely than the macro series. Similarly, if the prior is set stronger on the covariance of the macro series, then the state variables will fit the economic series over the yields.

There is an additional wrinkle in this problem, since $W_t$ and $U_t$ are not uncorrelated. Theory suggests that the errors in the consumption and CPI growth rates to be correlated with the errors in their respective drift processes to accommodate for the stochastic risk-premia. If they were uncorrelated, then there wouldn’t be any market prices of risk from
the \( \mu \)'s. Specifically, the covariance between \( W_t \) and \( U_t \) is represented by \( \Sigma_t \). The first row of all zero's imply that the drift rates are uncorrelated with respect to the pricing error. Only the error in the drift rate in consumption is correlated with the error of its drift rate \( \mu_t^c \), and similarly for CPI. Since the equations in (19)-(20) are correlated, they can be orthogonalized by writing them conditionally as

\[
\begin{align*}
y_t &= m_t + b\mu_t + W_t & W_t &\sim N_3(0, W) \\
\mu_t &= e + \tilde{d}_{t-1}\mu_{t-1} + j_{t-1}(y_{t-1} - m_{t-1}) + U_t & U_t &\sim N_2(0, \tilde{V}_{t-1})
\end{align*}
\]

where \( j_{t-1} = \Sigma_t' W^{-1} \), \( \tilde{d}_{t-1} = d - j_{t-1}b \) and \( \tilde{V}_{t-1} = V_{t-1} - j_{t-1}\Sigma_{t-1} \).

Now, the filtering problem can be re-written as

\[
\begin{align*}
\text{tions (21)-(22) form the observation equations. Note that only equation (21) is truly orthogonalized by writing them conditionally as }
\end{align*}
\]

\[
\begin{align*}
y_t &= m_t + b\mu_t + W_t & W_t &\sim N_3(0, W) \\
\mu_t &= e + \tilde{d}_{t-1}\mu_{t-1} + j_{t-1}(y_{t-1} - m_{t-1}) + U_t & U_t &\sim N_2(0, \tilde{V}_{t-1})
\end{align*}
\]

Now, draw from \( \mu_t | y_t, v_{t-1}, \Theta \) by using the Kalman filter via the forward-filtering backward sampling (FFBS) algorithm described by Carter and Kohn (1993). A slightly generalized version needed in this case is described in the appendix. Briefly, in order to simulate from this, first run the Kalman filter forward to get the moments of \( p(\mu_t | y_t, v_{t-1}, \Theta) \). Then sample the last state from \( \tilde{\mu}_T = p(\mu_T | y_T, v_{T-1}, \Theta) \) and sample backward through time \( \tilde{\mu}_t = p(\mu_t | \mu_{t+1}, y_t, v_{t-1}, \Theta) \), where \( y_T^0 = [y_T, \ldots, y_1] \). Then the samples \( \tilde{\mu}_1, \ldots, \tilde{\mu}_T \) are drawn jointly from \( p(\mu | y, v, \Theta) \), which makes the draws less autocorrelated and convergence is quick.

The next filtering problem is to draw the volatility states \( v_t | y_t, \mu_t, \Theta \). Assume the following notation:

\[
f = \left[ \begin{array}{c} \beta_0^c \\ 0 \end{array} \right], \ g = \left[ \begin{array}{cc} (1 - \beta_1^c) & 0 \\ 0 & (1 - \beta_1^c) \end{array} \right] \text{ and } E = \left[ \begin{array}{cc} \varepsilon_c^2 & 0 \\ 0 & \varepsilon_q^2 \end{array} \right]
\]

Now, the filtering problem can be re-written as

\[
\begin{align*}
\begin{bmatrix} \tilde{y}_t^{P(\tau_1)} \\ \vdots \\ \tilde{y}_t^{P(\tau_n)} \end{bmatrix} &= \begin{bmatrix} D(\tau_1)\beta_0^c + E(\tau_1)\beta_1^c + F(\tau_1)\beta_2^c + G(\tau_1)\beta_3^c \\ \vdots \\ D(\tau_n)\beta_0^c + E(\tau_n)\beta_1^c + F(\tau_n)\beta_2^c + G(\tau_n)\beta_3^c \end{bmatrix} + N_t^1 \quad N_t^1 \sim N(0, \varepsilon_c^2 I_{n\times n}) (21)
\end{align*}
\]

\[
\begin{align*}
\tilde{\mu}_{t+1} &= U_t & U_t &\sim N_2(0, \tilde{V}_t) \\
v_{t+1} &= f + g\tilde{v}_t + E_t & E_t &\sim N_2(0, E) (22)
\end{align*}
\]

where

\[
\begin{align*}
\tilde{y}_t^{P(\tau_j)} &= y_t^{P(\tau_j)} - A(\tau_j) - \left[ \begin{array}{c} B(\tau_j) \\ C(\tau_j) \end{array} \right] \tilde{\mu}_t \\
\tilde{\mu}_{t+1} &= \mu_{t+1} - \tilde{d}_t - j_t(y_t - m_t)\mu_t
\end{align*}
\]

Equations (21)-(23) form the state-space system for the volatility filtering problem. Equations (21)-(22) form the observation equations. Note that only equation (21) is truly
observable in the sense that we observe the yield curve. Equation (22) is observable in the sense that the volatility filter is computed conditional on the drift rates $\mu$. This filtering problem is highly non-linear and there are no known optimal way to filter the volatility. Volatility and volatility squared appears in the mean level of the yield curve and also enters in the volatility of the growth rates. The algorithm that is used here is a generalized version of the FFBS algorithm discussed in the filtration of the growth rates. This generalized method is jointly developed in Hore and McCulloch (2005) and is described in the appendix. Again, the strength of prior information will determine which of the observable series will contribute more to the filtering of the volatilities. If $\epsilon_\tau$ is very strong, then the volatilities will be drawn primarily from the yield curve. As the prior is relaxed, the growth rates will start to dominate and the volatilities will be drawn from the growth rates. A moderate level of $\epsilon_\tau$ will tend to draw the volatility from both series.

Thus far, filtering stochastic volatility has been restricted to the work of Jacquier, Rossi and Polson (1994) whose approach is drawing the states of volatility one at a time in a Metropolis setting, but the volatility draws are highly autocorrelated and convergence is slow. Shephard and Kim (1998) have a different approach of filtering stochastic volatility, but their approach does not correspond to the structural model here. Ahn et. al. investigate quadratic term structure models and use simulated method of moments to circumvent the filtering problem by using the reprojection method proposed by Gallant and Tauchen (1998). Gala and Veronesi use approximate maximum likelihood to estimate stochastic volatility. Our approach is flexible because it can take any functional form of volatility, and we draw the entire volatility as a block that reduces autocorrelation and convergence is fast.

This method filters the entire set of volatilities $v = \{v_1, \ldots, v_t\}$ jointly as a block as the posterior draw from observing both yield and growth rates with a given transition probability. This is a much more efficient algorithm than JPR in the sense that all the volatilities are drawn at once which reduces autocorrelation and convergence is quick. The drawback of this method is that it is a discrete filter and discretization may not be perfect if the true volatility lies somewhere in the middle of two points in the grid. To alleviate the problem a fine grid is chosen, but then this method becomes very costly computationally depending on how fine the grid is. Also, depending on the prior information, learning about the true value of the states can be very quick or slow depending on how informative the data is about the volatilities.

### 3.2 Posterior distribution of parameters

Having known the filtered volatility and growth rates, now the rest of the parameters can be obtained via the MH algorithm. The rest of the parameters are the regression, volatility, correlation and utility parameters. Closed-form posterior distributions are available without including the yield-curve. However, when the yield-curve is added, no closed form solutions
are available for the specific parameters and a likelihood based estimation technique is used in conjunction with the gibbs sampler. 

Regression Parameters: The posterior distribution of the regression parameters \((e, d)\) are discussed here. Exactly the same methodology applies to \((f, g)\) and are skipped for brevity. The full posterior distribution of the regression parameters conditional on the other parameters as well as the filtered states is

\[
p(e, d | \Theta_{-e,-d}, \mu, v, y) \propto p(e, d) p(\mu | v, e, d, \Theta_{-e,-d}) p(y^{P(\tau)} | e, d, v, \mu, \Theta_{-e,-d})
\]  

(24)

The first term on the right is the prior distribution on the regression parameters. A convenient prior choice is multivariate normal because it is conjugate with the first likelihood term, which is the second term on the right. The second term is the likelihood of observing the latent drift rates \(\mu\) conditional on the regression parameters, volatility and the rest of the parameter space. If the yield-curve is left out from the computation of the posterior, then these two densities jointly form the posterior distribution of the regression parameters. If the prior distribution is multivariate normal, then because of conjugacy rules, the posterior is also a multivariate normal with its mean centered at a point which is a linear combination between the OLS estimates and the prior mean. However, with the addition of the yield curve, no straightforward posterior density is available because the regression parameters enter the yield equation non-linearly through the factor loadings. As such, a metropolis step is necessary in order to sample from the full posterior. A random-walk metropolis is implemented in order to simulate from the posterior density (24).

The algorithm goes like the following. At iteration \(j\), draw a candidate \((e, d)^j\) from a certain proposal density based on the previous draw of the parameters \((e, d)^{j-1}\). In this case, consider a multivariate normal with a mean centered at the previous estimate. Now, accept this candidate as the posterior draw with probability \(\alpha((e, d)^j, (e, d)^{j-1})\) where

\[
\alpha((e, d)^j, (e, d)^{j-1}) = \min \left( \frac{p(e, d)^j p(\mu | v, (e, d)^j, \Theta_{-e,-d}) p(y^{P(\tau)} | (e, d)^j, v, \mu, \Theta_{-e,-d})}{p(e, d)^{j-1} p(\mu | v, (e, d)^{j-1}, \Theta_{-e,-d}) p(y^{P(\tau)} | (e, d)^{j-1}, v, \mu, \Theta_{-e,-d})} \right),
\]  

(25)

Essentially, propose a certain candidate for the posterior draw from a distribution that is “close” and choose to accept that as a posterior draw by performing a likelihood-ratio test. The algorithm will accept the new point for sure if the likelihood went up, otherwise it would accept the new point with probability as specified in (25). In essence, the new point will be accepted if the combined likelihood from both the growth rates and yield curve is greater than at the previous point. There is a possibility that the algorithm will accept a draw at a lower likelihood, but the lower likelihood will be discarded once another point with a higher likelihood is drawn from the proposal. The technical issues of convergence of such chains go back to the idea of time-reversibility, invariance, etc. and are discussed in detail in Tierney(1994) and Polson and Johannes(2004).
Volatility parameters: There are two sets of volatility parameters - the volatility of the volatility equations (23) and the volatility of the yield curve which we imposed to get the state-space system. The methodology for all of them is the same and only the volatility of the volatility of the consumption growth rate is discussed in detail. The posterior distribution for the parameter \( \epsilon^2_c \) is

\[
p(\epsilon^2_c | \Theta_{-\epsilon^2_c}, \mu, v, y) \propto p(\epsilon^2_c)p(v_c|\mu, \epsilon^2_c, \Theta_{-\epsilon^2_c})p(y^{P(r)}|v, \mu, \epsilon^2_c, \Theta_{-\epsilon^2_c})
\]  

(26)

With a prior conjugate probability distribution of inverted chi-squared, the posterior without the yield curve is also an inverted chi-squared. The posterior distribution for the volatility also depends on the prior and posterior regression parameters. However, if the prior distribution on the regression parameters is sufficiently diffuse, and if the sample size is big then the posterior draw is close to the OLS estimate of the volatility. After adding in the yield equation, a closed form solution is no longer available. In this case, make the proposal density for the \( j \)th draw

\[
\tilde{p}(\epsilon^2_c) \propto p(\epsilon^2_c)p(v_c|\mu, \epsilon^2_c, \Theta_{-\epsilon^2_c})
\]

(27)

and accept it with a probability generated by a posterior ratio test

\[
\alpha((\epsilon^2_c)^j, (\epsilon^2_c)^{j-1}) = \min \left( \frac{p(y^{P(r)}|(\epsilon^2_c)^j, v, \mu, \Theta_{(\epsilon^2_c)})}{p(y^{P(r)}|(\epsilon^2_c)^{j-1}, v, \mu, \Theta_{(\epsilon^2_c)})}, 1 \right).
\]

(28)

There is an exception for estimating \( \epsilon^2_c \), in which case the draw is a direct conjugate inverted chi-square.

Also, there is a bivariate variance-covariance matrix which determines the covariance structure of the bottom two observation equations in (19), i.e. the error in the observation of consumption and price level growth. The conjugate prior for variance covariance matrices is inverted Wishart, and a draw of a multivariate wishart is accepted as a block by using the same likelihood ration test as in (25).

Correlation parameters: The correlation coefficients which represent the correlation in the brownian motion terms between the drift rates and the consumption/CPI processes drive the market price of risk. The correlation matrix is hidden within \( \Sigma_t \) in the modified state-space system in equation (20), which is repeated here for convenience.

\[
\mu_t = e + \tilde{d}_{t-1}\mu_{t-1} + j_{t-1}(y_{t-1} - m_{t-1}) + U_t \quad U_t \sim N_2(0, \tilde{\Sigma}_{t-1})
\]

(29)

where \( j_{t-1} = \Sigma'_{t-1}W^{-1}, \tilde{d}_{t-1} = d - j_{t-1}b \) and \( \tilde{V}_{t-1} = V_{t-1} - j_{t-1}\Sigma_{t-1} \). The time subscript is dropped from \( \Sigma \) because the filtered volatility states are being conditioned on. However,
keeping the notation to $\Sigma$ is convenient because that is how the correlations enter the system. The estimation of these correlations is tricky because no closed-form solution exists even for the reduced model without the yield curve. The correlation enters through the regression coefficients $\tilde{d}, j$ and volatility $V$. The full posterior distribution for $\Sigma$ is

$$p(\Sigma|\Theta_{-\Sigma}, \mu, v, y) \propto p(\Sigma)p(\mu|\Sigma, \Theta_{-\Sigma}, v, y^c, y^o)p(y^{P(\tau)}|\Sigma, \Theta_{-\Sigma}, v, \mu)$$ \hspace{1cm} (30)$$

Even when the yield curve equation is dropped, no easy conjugate solution exists because the correlation enters through both the regression coefficients and the error term. However, the Metropolis algorithm is flexible enough to handle this situation. To draw the correlations, propose a draw from the prior $p(\Sigma)$. Take the prior distribution to be $U(-a, b)$, where $a$ and $b$ are chosen such that $\Sigma$ will still have positive eigenvalues. Then accept this draw with probability

$$\alpha(\Sigma^j, \Sigma^{j-1}) = \min\left(\frac{p(\mu|\Sigma^j, \Theta_{-\Sigma^j}, v, y^c, y^o)p(y^{P(\tau)}|\Sigma^j, \Theta_{-\Sigma^j}, v, \mu)}{p(\mu|\Sigma^{j-1}, \Theta_{-\Sigma^{j-1}}, v, y^c, y^o)p(y^{P(\tau)}|\Sigma^{j-1}, \Theta_{-\Sigma^{j-1}}, v, \mu)}, 1\right).$$ \hspace{1cm} (31)$$

**Utility parameters** The utility parameters - risk aversion ($\gamma$) and discount factor($\phi$) can only be estimated through the yield curve. Because of the imposition of a pricing error, evaluate the likelihood for the parameters over a set of values. Convert them into a probability distribution over those set of values, and then the utility parameters can be sampled from that probability distribution. The range of values (i.e. grid) is chosen such that it is consistent with both theory and empirical knowledge. The grid must also be chosen with certain flexibility such that the bounds are never binding.

### 4 Empirical Findings

**Data:** The data used for the empirical investigation comes from monthly US real consumption (non-durables and services) data (1959-2004) divided by population. Monthly price level data comes from International Financial Statistics, while yield curve data comes from the 1-5 year bond prices from the Fama-Bliss database in CRSP.

**Priors** Before going into the empirical discussion, it is important to discuss the set of priors that will be considered. As has been discussed earlier, prior information will be used heavily to determine the role of interplay between the yield curve and macroeconomic data. Three prior settings are chosen to efficiently demonstrate the interplay between the two series. The choice of priors focuses around two key parameters. The first one is $\epsilon_r$, which controls the fit to no-arbitrage bond pricing, and the other is the covariance matrix of the economic fundamentals. The three prior choices are
tight: A tight prior on the error of the yield curve $\epsilon$, centered around 3 basis points and a very loose prior is set on the economic series. It can be considered to be a reduced form setting that will do a good job in fitting yields without much economic content.

loose: A loose prior of $\epsilon$ centered around 20 basis points, which is a very loose prior for monthly data. However, a very strong prior is set on the economic series. This can be considered as the uni-directional setting of Ang and Piazzesi (2003) which essentially says that bonds are not very informative for determining the dynamics of macro variables.

joint: A joint prior is set such that the resulting set-up is a mixture of both yields and the macroeconomic data. This incorporates the setting of Ang, Dong and Piazzesi (2005) by allowing for bi-directional macro-finance link under a no-arbitrage setting.

The key point to realize is that for every specification a no-arbitrage model along with a bi-directional macro-finance link is considered. What separates one setting from another is which series is enforced to be more informative than the other by prior settings. This estimation technique produces the result as expected. The tight prior specification produces state-variables and parameters that fit the yield curve very well, but do not do a good job in fitting the macroeconomic series. Similarly, the loose prior specification fits the economic series well but do not do a good job in fitting the yields. The joint prior specification does an intermediate job in fitting the yields and the macro data. It turns out that the joint “mixture” setting has the most desirable feature when it comes to predicting future yields, but the tight setting does a much better job in matching the inflation risk-premia found in the literature. However, before discussing the implications of each model, a closer look at the underlying state variables of each model is presented. Ultimately, it is the flexibility in those state variables that will be used for prediction or checking stylized facts.

An examination of the fitted yields in Fig. 5 confirms the belief. The tight specification fits the yields very closely, while the loose specification does a much worse job. The joint specification finds an intermediate fit between the two as it considers information about state variables and parameters from both sources with roughly equal weights. Now, the relevant question is what does no-arbitrage imply about the underlying economic setting that cannot be revealed simply by an analysis of the macro factors themselves. For this answer, one needs to look into the latent growth rates and volatilities that are implied under different settings.

First, the tight setting implies consumption growth rates that are simply not observed in the economy. The top left graph in Fig. 1 shows the expected growth rates of consumption only derived from per capita consumption growth. In contrast, the top right figure shows the same growth rates under the tight prior specification. Obeying its prior settings, it does fit the bond yields well, however, as it does so it implies growth rates of consumption that are not observed in the economy. As the prior restriction is relaxed to the loose
setting, the resulting expected consumption growth matches the consumption series very closely (lower left graph of Fig. 1) but the resulting yield fit is very poor. Again, it reflects the prior setting which emphasized the macro dynamics and de-emphasized no-arbitrage bond prices. The intermediate joint setting does a moderate fit of the consumption data (lower right graph of Fig. 1), but the fit of yields is also modest. Therefore, one can use informative prior information about the level of mix between macro dynamics and no-arbitrage bond yields in order to filter state variables that can produce different kind of results. The two center graphs of Fig. 1 show what happens as prior information is relaxed. Clearly, the resulting expected growth rates of consumption shift into new levels from the tight specification getting closer to growth rates that correspond to what is observed in the economy at a cost of poorer fit of the yield curve.

For expected inflation the information from the bonds is not very different from what is seen in the economy except for certain periods of high recession, like the early eighties, nineties and two thousand. In these periods the information from the bond deviates sufficiently from the underlying economic model as can be seen from the first row of graphs in Fig. 2. Except for those periods, expected inflation does not change very significantly from the macro dynamics under a tight specification. Expected inflation in the joint setting looks like a compromise between the expected inflation from the pure macroeconomic analysis and that of the tight analysis (Fig. 2 center right graph).

So far, the implication of no-arbitrage is that it produces growth rates that may or may not look like the growth rates of the economic factors depending on the desired goodness of fit to the no-arbitrage model. The growth rates under different prior settings are not the only factors responsible for the different levels of yield curve fits. The state variables governing stochastic volatilities obtained under different settings are also different. Depending on which mix of no-arbitrage and macro factors is considered, determines what the volatility state variable will look like for both expected inflation and expected consumption growth. First the volatilities are estimated using the macroeconomic series alone, and the volatilities seem constant at about 2.5% yearly rate for consumption growth and 3% for expected inflation. However, when no-arbitrage is enforced these estimates start to change dramatically. Imposing the loose specification, the volatilities do not move significantly since the overall fit to the no-arbitrage bond yields in this setting is poor by construction. As such, it fails to capture the high level of volatility needed in underlying economic dynamics to capture bond yields. On the other extreme, when the tight prior setting is imposed, the no-arbitrage model prevails and information from the economic series is ignored. For both macro series, the time-series of volatilities implied from the tight setting that puts a heavy weight on the yield curve is greater by a factor of 10. Clearly, the volatilities of the underlying macro series as implied by no-arbitrage alone is not what is observed in the data. The “quasi”-reduced form model captures the high level of risk-premia that is observed in bonds, but not in a way that would correspond to the underlying macro dynamics. The most interesting part is the joint setting where a healthy mix of macro and no-arbitrage is
considered. As is previously conjectured, the resulting volatility lies in between the macro series and the reduced form setting. It pulls information from both sides and produces a compromising time-series of volatilities in the middle. Interestingly, the volatility in expected inflation falls dramatically in the mid-80s after Fed policy changes of aggressively fighting inflation. It removed from the minds of investors fears about unexpected inflation, thus reducing inflation risk-premia.

As has been seen in the literature (Diebold, Rudebusch and Auroba(2005), Ang and Piazzesi (2003)), adding macro information leads to a better explanatory power of the Nelson-Siegel type models. To assess the explanatory power of these models, model parameters and state-variables from each prior setting is used to predict future bond prices one year ahead. Due to the limited availability of yields, one year return on bonds are considered. The model-implied returns in-sample are compared against observed returns in the sample. Also, mean-squared errors of the prediction of next-period’s prices is reported in Table 3. The tight specification that fits the yield data very well (Fig. 5) does a poor job in predicting returns. In other words, a reduced form bond-pricing model that can fit the highly volatile bond market in the early-80’s does very poorly in predicting future bond prices. On the other hand, the joint specification with a healthy mix of no-arbitrage bond pricing along with macro information does a lot better job in predicting returns. Fig. 6 plots the two different specifications for two different maturities and finds clear evidence of superiority of the mixture model over the reduced form model in predicting returns. Comparing columns 1 and 3 of Table 2, one can see that the mixture model comfortably beats the reduced form model. Surprisingly, the loose specification that produces a bad fit of bond prices in Fig. 5 does a much better job in predicting return than the reduced form model as well.

However, as seen by the time-series of volatility estimates, the tight specification is the only one that can match the risk-premia from the existing literature. Wachter (2005) produced an inflation risk-premia of 48 basis points annually, but her estimate is not time-varying. Inflation risk-premia of a $\tau$-period bond here is the difference in the expected return between a nominal and a real bond in (15). The time-series of inflation risk-premia is presented in Fig. 7. Incidentally, due to the precipitous drop in the volatility of expected inflation in the mid-80s, inflation risk-premia also dropped significantly - a finding that cannot be inferred from her constant estimate. Essentially, as the Fed aggressively started fighting inflation, fears of unexpectedly high inflation went away from the market in the short term but has since come back to earlier levels. Inflation risk-premia between 30-70 basis points seem rather small for annual levels. The reason for this small number is due to the positive interaction between the real rate and expected inflation - a violation of the Fisher Hypothesis. If they were independent, a larger difference in their expected return would be observed. In this setting, a higher expected inflation gives a higher expected consumption growth rate which causes the real-rate to move along with the nominal rate resulting in a smaller estimate for inflation risk-premia than would be observed if they were
The parameter estimates are presented in Table 1 and graphs of the posterior distribution of some of the parameters are presented in Fig. 9 - Fig. 12. Notice that the posterior distribution of the parameters deviates significantly from their closed form posterior without the no-arbitrage condition. The violation of Fisher hypothesis can be verified by the posterior distribution of $\alpha_2$. The sign is negative, but the macro dynamics is with respect to $-\alpha_2$, which is positive. An interesting observation in this case is how does information from no-arbitrage effect the parameter values of macro dynamics. As can be seen from the posterior draw of the parameter in Fig. 9, the information from the bond tightens the posterior distribution of the parameter space in an interval within the support of the space under macro-dynamics solely. Notice that tightening of the parameter space takes place no matter what kind of blend of no-arbitrage and macro dynamics is considered. But, as expected, the tightening for different prior specifications is in different intervals. For example, the autoregressive parameter for the inflation process, $1 - \alpha_i$, has a much wider support under macro dynamics, but it tightens considerable when information from the yield curve is introduced. It tightens at different locations under different settings, and is more autoregressive as no-arbitrage gets stronger. Sometimes, a parameter may significantly deviate from the original support. For example, $\epsilon_c$ (Fig. 11) and $\epsilon_q$ move significantly away from its support under macroeconomic dynamics. This is not surprising since the time-series of volatility estimates are much greater and more volatile than its macro counterpart. In this instance a good proposal is hard to find and different priors should be considered for a good fit. Another parameter of interest is $\alpha_1$ (Fig. 13) which controls the effect of the real rate on the expected change in the nominal rate. In the underlying macro dynamics it is the feedback parameter of expected inflation in expected growth of consumption. Under the macro-dynamics, the posterior probability that the parameter is greater than zero is significant. In fact, its 95% posterior probability band contains zero. As such, there is a strong possibility in only judging from the macro dynamics that the parameter is insignificant. However, when the no-arbitrage setting is implied the parameter becomes considerably negative implying a positive effect of the real-rate on the expected change in the nominal rate. Again, the no-arbitrage restriction strongly identifies a parameter whose effect may look ambiguous in macro dynamics.

The risk-aversion coefficient is also very different under different levels of fit with the yield curve. Under the tight specification the posterior distribution of risk-aversion is much higher than the other two specifications that do not fit the yield data well. The risk-aversion parameter computed from the loose specification is very small, but that is because it is incapable of fitting the yield curve. We know from standard Euler equation calibration of the risk-free rate that a small risk-free rate would be delivered by a small risk-aversion parameter. Plugging in the high risk-aversion parameter from matching equity premium would generate risk-free rates that are much higher than observed. In fact, trying to calibrate small risk-free rates with high risk-aversion parameter generates inadmissible
discount rates, which is the famous risk-free rate puzzle. An out of sample test is considered where the two settings \textit{joint} and \textit{loose} that generate small risk-aversion parameters are fitted to the 30 day risk-free rate. The short-rate specification of (3) is used to compute the model implied 30-day risk-free rate. The assumption is for $\tau = 1$ (i.e. $dt = 1/12$), the coefficients $A \cdots G$ are very close to zero and hence the bond pricing equation (11) collapses to the short rate. The result of this out-of-sample study is produced in Fig. 8. Clearly the \textit{loose} and \textit{joint} settings produce risk-free rates that are much close to what we observe in the data. The \textit{tight} specification produces short-rates that are way off and are not reported.

5 Conclusion

An utility-based term-structure model is presented here in the presence of a nominal pricing kernel where the risk-premia is also endogenized. The stochastic volatility of expected consumption growth and expected inflation create time-varying risk-premia in a Quadratic-Gaussian model of the bond prices.

The model is estimated using a Bayesian methodology that considers a bi-directional influence - macroeconomic dynamics affecting the yield-curve, and vice versa. Three classes of models are considered, where each class corresponds to a prior specification that dictates the level of mixture between the yield-curve and macroeconomic dynamics. The class of prior specification that corresponds to a stronger influence of the yield curve does a good job in fitting the yield curve, but fails in prediction, does not correspond to the states of the underlying macro dynamics and does not fit the risk-free rate out of sample. Similarly, the specification that suggests stronger influence of macroeconomic dynamics and weak on the yield curve generates bad fit with the yield curve although it fits the macro dynamics perfectly. An in-between mixture of macro-dynamics and no-arbitrage bond prices does much better in prediction in-sample. The latter two also fit the 30-day risk-free rate because of low risk-aversion parameter obtained from those settings. However, the \textit{tight} specification matches the risk-premia in the literature very well.

A wide class of models can be considered under this Bayesian methodology trying to build a bridge between macroeconomic dynamics and no-arbitrage asset pricing. It is left upto future work to deal with more exotic models of macro-finance with jumps, and incorporating derivatives.
References


6 Appendix

**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

\[ M^n_t = e^{-\phi t} u'(c_t) = e^{-\phi t} c_t^{1-\gamma} \]

Applying Ito’s Lemma to \( M^n_t \) and using (1)-(2) we get

\[
\frac{dM^n_t}{M^n_t} = \left[ -\gamma \mu^c_t - \mu^q_q - \phi + \frac{\gamma(1+\gamma)\sigma_c^2}{2} + \sigma^2 \right] dt
- \gamma \sigma_c dW_c(t) - \sigma_q dW_q(t)
\]

Applying \( E_t[\frac{dM^n_t}{M^n_t}] = -r^n_t dt \) proves the result.

The real interest rate can be obtained by applying the same technique to the real pricing kernel \( M_t = e^{-\phi t} c_t^{-\gamma} \).

**Proof of Proposition (2.2):**
Given the nominal short rate process in (3), the price of a \( \tau \) period bond, \( P^n(\tau; \Theta) \), where \( \Theta \) is the entire parameter set, is

\[
P^n(\tau; \Theta) = E^Q \left[ \int_0^\tau \frac{1}{r^n_s(\Theta)} ds \right]
\]

Because of the differential form of the nominal rate process in (8), \( P^n \) must be a function of \( P^n(\mu^c_t, \mu^q_t, \vartheta^c_t, \vartheta^q_t, \gamma) \). At the same time, no arbitrage implies that \( P^n \) satisfies a stochastic differential equation (SDE) of the form (10). An application of Ito’s Lemma leads to

\[
\begin{align*}
\frac{1}{2} & \left[ P^n_{\mu^c \mu^c} c_t^{2} + P^n_{\mu^q \mu^q} c_q^{2} \right] dt + \left[ P^n_{\mu^c \vartheta^c} c_t^{2} + P^n_{\mu^q \vartheta^q} c_q^{2} \right] dt \\
& = \left[ \rho_1 \gamma \sigma_c P^n_{\mu^c \vartheta^c} c_t^{2} - \rho_2 \sigma_q P^n_{\mu^q \vartheta^q} c_q^{2} \right] dt
\end{align*}
\]

the solution of which is given by (11). After substituting (11) into the above PDE, we can verify the quadratic form and obtain the coefficients which solve a system of ordinary
differential equations given by

\[ A'(\tau) = \alpha_0^c B(\tau) + \alpha_0^q C(\tau) + \beta_0^c D(\tau) + \beta_0^q E(\tau) + \epsilon_c^2 F(\tau) + \frac{1}{2} \epsilon_q^2 D(\tau)^2 + \epsilon_q^2 G(\tau) + \frac{1}{2} \epsilon_q^2 E(\tau)^2 - k \]

\[ B'(\tau) = -\alpha_1^c B(\tau) - \alpha_1^q C(\tau) - \gamma \]

\[ C'(\tau) = -\alpha_2^c B(\tau) - \alpha_2^q C(\tau) - 1 \]

\[ D'(\tau) = 2\beta_0^c F(\tau) - \beta_1^c D(\tau) + 2D(\tau)F(\tau)\epsilon_c^2 + \gamma \sigma_c \rho_1 B(\tau) \]

\[ E'(\tau) = 2\beta_0^q G(\tau) - \beta_1^q E(\tau) + 2E(\tau)G(\tau)\epsilon_q^2 + \sigma_q \rho_2 C(\tau) \]

\[ F'(\tau) = -2\beta_1^c F(\tau) + \frac{1}{2} B(\tau)^2 + 2F(\tau)^2\epsilon_c^2 \]

\[ G'(\tau) = -2\beta_1^q G(\tau) + \frac{1}{2} C(\tau)^2 + 2G(\tau)^2\epsilon_q^2 \]

with initial condition \( A(0) = B(0) = C(0) = D(0) = E(0) = 0 \). 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 \( A \).

**FFBS algorithm with correlation in the error terms:**

Assume the following filtering problem

\[ X_t = F_t' \theta_t + \alpha_t + v_t \]

\[ \theta_{t+1} = G_t \theta_t + \gamma_t + w_t \] (32)

where the covariance structure is

\[ \begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim N\left( 0, \begin{pmatrix} V_t & \Sigma_t \\ \Sigma_t' & W_t \end{pmatrix} \right) \]

Once \( X_t \) is observed, then a peak at \( \theta_t \) is like observing \( v_t \)

\[ v_t = X_t - F_t' \theta_t - \alpha_t \] (33)

Now, the conditional distribution of \( w_t | v_t \) is

\[ w_t | v_t \sim N(\Sigma_t' V_t^{-1} (X_t - F_t' \theta_t - \alpha_t), W_t - \Sigma_t V_t^{-1} \Sigma_t') \] (34)

Now, (32) can be rewritten as

\[ X_t = F_t' \theta_t + \alpha_t + v_t \]

\[ \theta_{t+1} = H_t \theta_t + K_t (X_t - \alpha_t) + \gamma_t + z_t \] (35)

\[ 25 \]
where

\[ H_t = G_t - \Sigma_t' V_t^{-1} F'_t \]
\[ K_t = \Sigma_t V_t^{-1} \]

The error terms \( v_t \) and \( z_t \) are distributed independently

\[ v_t \sim N(0, V_t) \]
\[ z_t \sim N(0, Z_t) \]

where \( Z_t = W_t - \Sigma_t' V_t^{-1} \Sigma_t \).

Rewrite (35) as

\[ X_t = F_0 t + v_t \]
\[ \theta_t = H_{t-1} \theta_{t-1} + K_{t-1} (X_{t-1} - \alpha_{t-1}) + \gamma_{t-1} + z_{t-1} \] (36)

where \( v_t \) and \( z_{t-1} \) are still independent. Define \( D_t = (X_1, X_2, ..., X_t) \) to be the entire set of observation up to time \( t \). Let’s start off with a prior distribution of \( \theta_{t-1} | D_{t-1} \sim N(m_{t-1}, C_{t-1}) \). From that prior information, one needs to come up with \( \theta_t | D_t \sim N(m_t, C_t) \) after observing \( X_t \). Since (36) is jointly normal, all one needs to compute are the first two

\[ E(\theta_t | D_{t-1}) = H_{t-1} m_{t-1} + K_{t-1} (X_{t-1} - \alpha_{t-1}) + \gamma_{t-1} \equiv a_t \]
\[ V(\theta_t | D_{t-1}) = H_{t-1} C_{t-1} H_{t-1}^t + Z_{t-1} \equiv R_t \] (37)

Similarly,

\[ E(X_t | D_{t-1}) = F'_t a_t + \alpha_t \equiv f_t \]
\[ V(X_t | D_{t-1}) = F'_t R_t F_t + V_t \equiv Q_t \] (38)

and the covariance is

\[ \text{Cov}(\theta_t, X_t | D_{t-1}) = E(\theta_t \theta_t' F_t) = R_t F_t \] (39)

Now the conditional distribution of \( \theta_t | D_{t-1}, X_t \equiv \theta_t | D_t \) can be written as

\[ \theta_t | D_t \sim N(m_t, C_t) \] (40)
\[ m_t = a_t + A_t (X_t - f_t) \]
\[ C_t = R_t - A_t Q_t A_t' \]
\[ A_t = R_t F_t Q_t' \] (41)
Go forward in this manner and once $\theta_T|D_T$ is computed, draw the last state $\theta_T$ immediately. Then reverse the chain backwards drawing $\theta_t|\theta_{t+1}, D_t$. All the moments of the backward sampler is computed from the forward filtering.

$$p(\theta_t|\theta_{t+1}, D_t) \propto p(\theta_t|D_t)p(\theta_{t+1}|\theta_t, X_t)$$ (42)

Again, since everything is conditional normal, just compute the first two moments of the distribution

$$\theta_t|\theta_{t+1}, D_t \sim N(m_t + B_t(\theta_{t+1} - a_{t+1}), C_t - B_t R_{t+1} B'_t)$$

$$B_t = C_t(H'_t + F_t K'_t) R_{t+1}^{-1}$$ (44)

Generalized FFBS algorithm to filter volatility

A univariate version of the volatility state-space system in (21)-(23) is described here.

$$X_{t+1}^1 = c_1 v_{t+1} + c_2 v_{t+1}^2 + k Z_0$$ (45)

$$X_{t+1}^2 = v_{t+1} Z_1$$ (46)

$$v_{t+1} = \alpha + \beta v_t + c Z_2$$ (47)

where $Z_0$, $Z_1$ and $Z_2$ are $N(0,1)$ and independent of each other.

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 fine enough to go through the set of possibilities accurately and has to pick the bounds carefully such that the full realm of possibilities are covered. Notice, that every $y_i$ needs to be positive.

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, $p(v_1|X_1^1 = a, X_2^2 = b)$.

Notice that the conditional distribution of $v_1$ is given by the state equation (47)

$$p(v_1 = y_j|v_0 = y_k) \sim N(\alpha + \beta y_k, c\sqrt{y_k})$$ (48)

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

$$p(v_1 = y_j, v_0 = y_k) = p(v_1 = y_j|v_0 = y_k)p(v_0 = y_k)$$ (49)

The first term on the right is the transitional probability from state $y_k$ to state $y_j$ that one can obtain from (47). The second equation is the prior probability distribution which is given.
In order to get the posterior distribution of \( p(v_1|X_1^1 = a, X_1^2 = b) \), first compute the full joint distribution.

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

(50)

The first two terms on the right can be computed from (45)-(46). 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

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

(51)

The numerator is determined by (50). The denominator is a constant conditional on \( (X_1^1 = a, X_1^2 = b) \). Hence, the above can be written as

\[
p(v_1 = y_j, v_0 = y_k | X_1^1 = a, X_1^2 = b) \propto p(X_1^1 = a, X_1^2 = b, v_1 = y_j, v_0 = y_k)
\]

(52)

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

\[
p(v_1 = y_j | X_1^1 = a, X_1^2 = b) = \sum_{v_0 \in Y} p(v_1 = y_j, v_0 = y_k | X_1^1 = a, X_1^2 = b)
\]

(53)

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 \( v_1 \).

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

(54)

The first term on the numerator (denominator) is the transition probability in (47) while the second term is the obtained from the forward filtering (53).
Table 1: Posterior means of parameter estimates along with the 2.5% and 97.5% percentile interval that are obtained from the Markov chain generated from the MCMC sampling. Each Gibbs sampler is run for 15000 iterations and the first 10000 draws are discarded. The estimates presented here are from the next 5000 draws. “No Bonds” represents parameters estimates without the yield curve. “With Bonds” represents parameters estimated with the yield data of 1-5 years. Parameters are reported in the form that they are estimated and not in the original form in the structural model. The tight model is with pricing error 3.5bp and very little mixture of macro variables, the loose prior with 20bp and heavy interaction with macro variables and the joint prior with 6.5 basis points and an in-between mixture with macro variables.

<table>
<thead>
<tr>
<th></th>
<th>No Bonds</th>
<th>Tight</th>
<th>Loose</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.0035</td>
<td>0.0011</td>
<td>0.0032</td>
<td>0.0029</td>
</tr>
<tr>
<td>(0.0027,0.0044)</td>
<td>(0.0011,0.0012)</td>
<td>(0.0031,0.0033)</td>
<td>(0.0025,0.0033)</td>
<td></td>
</tr>
<tr>
<td>$1 - \alpha_1^T$</td>
<td>-0.4863</td>
<td>-0.4700</td>
<td>-0.4847</td>
<td>-0.4812</td>
</tr>
<tr>
<td>(-0.6521,-0.2736)</td>
<td>(-0.4715,-0.4690)</td>
<td>(-0.4874,-0.4828)</td>
<td>(-0.4840,-0.4782)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_2^T$</td>
<td>-0.2531</td>
<td>-0.2547</td>
<td>-0.2500</td>
<td>-0.2519</td>
</tr>
<tr>
<td>(-0.4646,-0.0602)</td>
<td>(-0.2550,-0.2543)</td>
<td>(-0.2509,-0.2496)</td>
<td>(-0.2523,-0.2512)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_0^L$</td>
<td>0.0005</td>
<td>0.0026</td>
<td>0.0011</td>
<td>0.0014</td>
</tr>
<tr>
<td>(0.0002,0.0017)</td>
<td>(0.0026,0.0027)</td>
<td>(0.0010,0.0011)</td>
<td>(0.0013,0.0015)</td>
<td></td>
</tr>
<tr>
<td>$1 - \alpha_1^L$</td>
<td>-0.1019</td>
<td>-0.1297</td>
<td>-0.0981</td>
<td>-0.0984</td>
</tr>
<tr>
<td>(-0.5610,0.1577)</td>
<td>(-0.1300,-0.1291)</td>
<td>(-0.0990,-0.0971)</td>
<td>(-0.0996,-0.0975)</td>
<td></td>
</tr>
<tr>
<td>$\beta_0^T$</td>
<td>0.9009</td>
<td>0.9226</td>
<td>0.9018</td>
<td>0.9029</td>
</tr>
<tr>
<td>(0.7825,0.9705)</td>
<td>(0.9220,0.9228)</td>
<td>(0.9013,0.9020)</td>
<td>(0.9022,0.9032)</td>
<td></td>
</tr>
<tr>
<td>$\beta_0^L$</td>
<td>0.0004</td>
<td>0.0048</td>
<td>0.0005</td>
<td>0.0017</td>
</tr>
<tr>
<td>(0.0002,0.0005)</td>
<td>(0.0047,0.0049)</td>
<td>(0.0003,0.0007)</td>
<td>(0.0005,0.0032)</td>
<td></td>
</tr>
<tr>
<td>$1 - \beta_1^T$</td>
<td>0.8551</td>
<td>0.8850</td>
<td>0.8560</td>
<td>0.8506</td>
</tr>
<tr>
<td>(0.8228,0.8867)</td>
<td>(0.8822,0.8864)</td>
<td>(0.8516,0.8586)</td>
<td>(0.8461,0.8598)</td>
<td></td>
</tr>
<tr>
<td>$\beta_0^L$</td>
<td>0.0002</td>
<td>0.0013</td>
<td>0.0003</td>
<td>0.0007</td>
</tr>
<tr>
<td>(0.0001,0.0003)</td>
<td>(0.0013,0.0013)</td>
<td>(0.0001,0.0004)</td>
<td>(0.0006,0.0007)</td>
<td></td>
</tr>
<tr>
<td>$1 - \beta_1^L$</td>
<td>0.8587</td>
<td>0.8632</td>
<td>0.8528</td>
<td>0.8588</td>
</tr>
<tr>
<td>(0.8244,0.8893)</td>
<td>(0.8620,0.8632)</td>
<td>(0.8511,0.8570)</td>
<td>(0.8585,0.8594)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^T$</td>
<td>0.0014</td>
<td>0.0088</td>
<td>0.0019</td>
<td>0.0016</td>
</tr>
<tr>
<td>(0.0013,0.0016)</td>
<td>(0.0078,0.0092)</td>
<td>(0.0018,0.0020)</td>
<td>(0.0015,0.0017)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^L$</td>
<td>0.0009</td>
<td>0.0081</td>
<td>0.0043</td>
<td>0.0053</td>
</tr>
<tr>
<td>(0.0008,0.0010)</td>
<td>(0.0080,0.0082)</td>
<td>(0.0047,0.0050)</td>
<td>(0.0050,0.0054)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^{T^2}$</td>
<td>0.0003</td>
<td>0.0089</td>
<td>0.0004</td>
<td>0.0007</td>
</tr>
<tr>
<td>(0.0007,0.0315)</td>
<td>(0.0080,0.0094)</td>
<td>(0.0003,0.0004)</td>
<td>(0.0005,0.0009)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^{L^2}$</td>
<td>0.0028</td>
<td>0.0041</td>
<td>0.0002</td>
<td>0.0017</td>
</tr>
<tr>
<td>(0.0026,0.0032)</td>
<td>(0.0038,0.0045)</td>
<td>(0.0002,0.0002)</td>
<td>(0.0016,0.0020)</td>
<td></td>
</tr>
<tr>
<td>$\rho^T$</td>
<td>-0.0174</td>
<td>-0.5382</td>
<td>-0.1332</td>
<td>-0.1067</td>
</tr>
<tr>
<td>(-0.0556,-0.0047)</td>
<td>(-0.8334,-0.3501)</td>
<td>(-0.1770,-0.1455)</td>
<td>(-0.2752,-0.0467)</td>
<td></td>
</tr>
<tr>
<td>$\rho^L$</td>
<td>0.0334</td>
<td>0.0046</td>
<td>0.0525</td>
<td>0.0206</td>
</tr>
<tr>
<td>(0.0140,0.1161)</td>
<td>(0.0043,0.0048)</td>
<td>(0.0059,0.1526)</td>
<td>(0.0012,0.0045)</td>
<td></td>
</tr>
<tr>
<td>$\rho_2^T$</td>
<td>0.0502</td>
<td>0.0811</td>
<td>0.0159</td>
<td>0.0355</td>
</tr>
<tr>
<td>(0.0046,0.1506)</td>
<td>(0.0542,0.0.09)</td>
<td>(0.0007,0.0736)</td>
<td>(0.0062,0.1245)</td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-2.709</td>
<td>0.0882</td>
<td>0.6549</td>
<td></td>
</tr>
<tr>
<td>(2.6770,2.753)</td>
<td>(0.0267,0.3378)</td>
<td>(0.4449,1.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Annual estimates of market prices of risk from the model in (12). Given the distribution of parameter values in Table 1, the market prices of risk for $\theta^c_t$ and $\theta^q_t$ are computed for each bond of maturity $\tau$. The median level, as well as the 2.5% and 97.5% percentiles, of the values are reported. Only up to four significant digits are presented.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>tight</th>
<th>loose</th>
<th>joint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta^c$</td>
<td>$\theta^q$</td>
<td>$\theta^q$</td>
</tr>
<tr>
<td>2</td>
<td>0.0010</td>
<td>0.0040</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0005,0.0060)</td>
<td>(0.0036,0.0088)</td>
<td>(0.0000,0.0018)</td>
</tr>
<tr>
<td>3</td>
<td>0.0008</td>
<td>0.0045</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0004,0.0054)</td>
<td>(0.0032,0.0340)</td>
<td>(0.0000,0.0019)</td>
</tr>
<tr>
<td>4</td>
<td>0.0008</td>
<td>0.0049</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0004,0.0052)</td>
<td>(0.0045,0.0364)</td>
<td>(0.0000,0.0020)</td>
</tr>
<tr>
<td>5</td>
<td>0.0007</td>
<td>0.0050</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0003,0.00050)</td>
<td>(0.0046,0.0384)</td>
<td>(0.0000,0.0021)</td>
</tr>
</tbody>
</table>

Table 3: Mean-squared error of prediction of expected return of fitted model vs. actual return. The expression in (12) is estimated and tested against one-year ahead yields as predicted by the model. In this case, only the median values of the parameter and state-variable estimations are used. Errors are in units of basis points.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>tight</th>
<th>loose</th>
<th>joint</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>78.5819</td>
<td>41.2124</td>
<td>33.0882</td>
</tr>
<tr>
<td>3</td>
<td>80.3594</td>
<td>51.2917</td>
<td>43.6632</td>
</tr>
<tr>
<td>4</td>
<td>83.8440</td>
<td>60.1562</td>
<td>53.1105</td>
</tr>
<tr>
<td>5</td>
<td>87.7185</td>
<td>67.2217</td>
<td>60.6111</td>
</tr>
</tbody>
</table>
Figure 1: Filtered Expected Monthly Consumption Growth Rate $\mu_t^c$
Figure 2: Filtered Expected Monthly Inflation $\mu_t^q$
Figure 3: Filtered Monthly Volatility of Expected Inflation
Figure 4: Filtered Monthly Volatility in Expected Consumption Growth Rate
Figure 5: Fitted Yields
Figure 6: Bond return
Figure 7: Annualized Inflation Risk-Premia
Figure 8: 30 day Risk-free rate (Monthly)
Figure 9: Interaction between expected inflation and change in expected consumption growth
Figure 10: Autoregressive parameter of expected inflation
Figure 11: Standard deviation of the volatility in expected consumption growth
Figure 12: Risk Aversion
Figure 13: Interaction between expected consumption growth and change in expected inflation