Abstract

This paper assesses the relative importance of long-run risk and valuation risk in accounting for the behavior of stock returns and nominal bond yields. I do so by estimating a consumption-based asset-pricing model that allows for both types of risk. I argue that valuation risk, modeled as persistent shocks to agents’ discount rates, plays a key role in accounting for the salient properties of the nominal yield curve. I also show that valuation risks are more correlated with statistical affine factors than long-run risks. I find that the valuation risks enter into the standard affine term structure model in a statistically significant manner, playing a particular role in accounting for movements in the long end of the yield curve.

1 Introduction

Standard consumption-based asset pricing models struggle to reconcile high and volatile equity premiums, defined as the difference between stock returns and short-term interest rate, with relatively smooth macroeconomic fundamentals. Recently, two classes of structural models have shown potential to resolve this matter. The long-run risk model, pioneered by [Bansal and Yaron (2004)], emphasizes the importance of uncertainty regarding the supply side of the economy in the form of persistent shocks to the consumption growth process. The valuation risk model, proposed by [Albuquerque et al. (2015)], emphasizes demand side shocks, modeled as persistent variations in agents’ time discount factor. When coupled with the assumption that agents prefer early resolution of risk, either source of risk can generate a high equity premium. However, these two models have potentially very different implications for the key characteristics of the nominal yield curve. This paper uses data on nominal
yields to assess the relative importance of valuation and long run risk shocks.

I proceed in two steps. First, I estimate a consumption-based asset pricing model that allows for both long-run and valuation risks. Second, I relate the estimated long-run and valuation risk shocks to statistical factors from a standard affine term structure model. I find that both risks are important in determining stock returns. Valuation risks play a critical role in accounting for the key properties of the nominal yield curve, including its upward-sloping nature and the variability of the long-term bond yields.

In my model, a representative agent makes consumption and savings decision in an endowment economy. The agent prefers early resolution of uncertainty and faces persistent variation in her time discount rate. Both consumption growth and dividend growth are driven by a small and persistent long-run risk component. All shocks to economic fundamentals feature time-varying stochastic volatility, following an autoregressive gamma process. I estimate the model’s parameters and the unobserved state variables by Bayesian likelihood-based Markov Chain Monte Carlo (MCMC) methods using data on consumption growth, dividend growth, inflation, stock returns, the short-term real interest rate and nominal bond yields. To improve estimation efficiency I implement a recently developed technique, the particle Gibbs sampler, proposed by Andrieu, Doucet, and Holenstein (2010), when sampling unobserved stochastic volatility states.

My key results from this economic model are as follows. First, the estimated model fits both stock prices and nominal bond yields reasonably well. Second, both long-run risk and valuation risk are important determinants of stock price movement, each accounting for approximately 40% of total variance of stock returns. Third, the model generates an upward-sloping nominal yield curve. According to my estimates, valuation risks account for roughly 60% and 70% of the variation in short-term and long-term nominal bond yields respectively. In contrast long-run risks account for roughly 10% of short and long term bond yields. Fourth, my estimated model implies that absent valuation risks, the nominal yield curve would be downward-sloping.

The intuition behind those results are as follows. Valuation risk and long run risk have different implications for shape of the real yield curve. The long-run risk model generally implies a downward-sloping real yield curve, because agents are concerned that consumption growth may be drastically lower for an extended period of time in some future state of the world. Such agents would like to buy real bonds that promise certain payout in all states of the world as a hedge. The longer the maturity of the bond, the more insurance it offers and the higher is its price. In contrast, the valuation risk model implies an upward-sloping real yield curve. This is because uncertainty about the time discount rate is increasing in the asset’s maturity, and therefore short-term bonds provide more insurance than their long-term counterparts. Since the nominal yield curve is upward-sloping in the data, the long-run risk model needs a sharply negative correlation between consumption growth and inflation to
justify both a downward-sloping real yield curve and an upward-sloping nominal yield curve. However, such sharp a negative correlation is not present in the data. Valuation risk model, in contrast, already implies an upward sloping real yield curve. So it only needs a weaker correlation between inflation and consumption growth to justify an upward-sloping nominal yield curve.

In the second part of the paper I explore the connection between economic factors estimated using my consumption-based model and the statistical factors extracted using a canonical Affine Term Structure Model (ATSM). An important advantage of the ATSM approach is that it requires less restrictive assumptions than my economic model. The downside is that the resulting factors have no formal economic interpretation. I show that valuation risk factors are more correlated with affine factors than long-run risk factors. Moreover, when I add valuation risks as an independent observable factor in an augmented ATSM, the estimated model attributes a large fraction of the variation in long-term bond yields to the valuation risk factor. Variance decomposition exercises suggest that valuation risk shocks account for a significant fraction of the variance of yields on long-term bonds. My results complement those of Ang and Piazzesi (2003) who find, using similar methods, that economic activity and inflation are important factors in explaining the variation in yields on short term bonds.

My paper is related to the literature as follows. The equity premium puzzle has been a challenge for economists. Ever since Bansal and Yaron (2004), many studies have used long-run risk models to explain asset pricing moments, see Bansal, Kiku, and Yaron (2012), Bansal and Shaliastovich (2013), Beeler and Campbell (2012), and Schorfheide, Song, and Yaron (2016). Following these authors, my consumption-based model features a persistent component in consumption growth and persistent stochastic volatilities. I follow Albuquerque et al. (2016) in allowing persistent and stochastic variation in the agent’s time discount rate. To highlight the role of valuation risk in explaining the equity premium, Albuquerque et al. (2015) abstract from the long-run risk component and estimate model parameters by GMM using stock returns and short-term real interest rate data. My work complements that paper by also incorporating nominal yields data.

My analysis is related to several papers that study the implications of long-run risk models for bond returns. Bansal and Shaliastovich (2013) study nominal bond pricing in a long-run risk model. Their model generates an upward-sloping nominal yield curve and downward-sloping real yield curve, by relying on a negative correlation between the persistent components of consumption growth and inflation. Song (2014) studies a modified version of the long-run risk model with exogenous regime-switching in monetary policy aggressiveness and regime-switching in the correlation between consumption growth and inflation. He finds that this flexible model is able to fit both stock prices and the nominal yield curve. My model offers a relatively parsimonious alternative that highlights the role of valuation risk. Creal and Wu (2016) nest an external-habit-style model as in Abel (1990) and Campbell and
Cochrane (1999) into the long-run risk framework and study nominal bond pricing. Without an external habit, Creal and Wu’s long-run risk model generates a downward-sloping nominal yield curve, and they show that adding an external habit can reverse this result. Albuquerque, Eichenbaum, Papanikolaou, and Rebelo (2015) argue that long-run risk models with external habit imply a counterfactually high correlation between stock returns and consumption growth. My model accounts for the nominal yield curve in a way that is consistent with the fact that consumption growth displays very little correlation with stock returns. Schorfheide, Song, and Yaron (2016) also nest long-run risk and valuation risk in their analysis. They find that valuation risk shocks improve the model’s fit with respect to the short-term real interest rate, however they do not study nominal bond yields per se.

From a methodological perspective, my paper relates to recent works that estimate consumption-based models using Bayesian likelihood-based method. Schorfheide, Song, and Yaron (2016) and Song (2014) directly evaluate the likelihood function of their model using the particle filter. Creal and Wu (2016) carry out model inference using a two-step approach. They first apply a Gibbs sampling procedure to estimate dynamics of economic fundamentals using inflation and consumption growth data. Taking parameter and state variable estimates as given, they then estimate representative agents’ preference parameters using asset price data. I adopt elements of their method with modifications. Similar to Song (2014) and Schorfheide et al. (2016), I simultaneously estimate the process for economic fundamentals and asset prices. Instead of using the particle filter, I apply a Gibbs sampling procedure that exploits the block structure of my non-linear state space model. One of the Gibbs sampling steps involves drawing sequences of stochastic volatility states. In that step, I follow Creal and Wu (2016) to use the particle Gibbs sampler, but I apply it to both economic fundamentals and asset price data. My paper is also related to the stochastic volatility literature in macroeconomics, for example Justiniano and Primiceri (2008). They exploit the block structure between stochastic volatilities and other parts of their model, so that a simpler sampling method proposed by Kim, Shephard, and Chib (1998) can be applied. In my application, the inclusion of asset prices makes it necessary to use the particle Gibbs sampler.

The rest of my paper is organized as follows. Section II presents the economic model nesting both long-run risk and valuation risk. Section III explores the connection between economic factors extracted from consumption-model to statistic factors from ATSM. Section IV concludes.

2 Consumption-based Asset Pricing Model

2.1 Economic Environment

In this section, I describe a representative-agent endowment economy featuring both long-run and valuation risks. The representative agent has recursive preferences as in Epstein and Zin (1991) and Weil (1989). When making consumption and savings decision, the agent
cares about current consumption $C_t$ and the certainty equivalent of future utility $U_{t+1}^*$. The representative agent’s life-time utility as a function of wealth $W_t$ is given by,

$$U_t(W_t) = \max_{C_t} \left[ \lambda_t \cdot C_t^{1-\frac{1}{\psi}} + \delta \cdot (U_{t+1}^*)^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\psi}}$$

subject to the budget constraint,

$$W_{t+1} = R_{c,t+1} \cdot (W_t - C_t)$$

where $R_{c,t+1}$ denotes the gross return on an asset paying aggregate consumption, which the agent takes as given. By definition $U_{t+1}^*$ is the sure value of $t+1$ lifetime utility such that

$$(U_{t+1}^*)^{1-\gamma} = E_t (U_{t+1}^{1-\gamma}).$$

The parameter $\delta$ is the agent’s time discount rate in the steady state; parameters $\gamma$ and $\psi$ control the agent’s relative risk aversion and the elasticity of inter-temporal substitution, respectively. In the special case where $\gamma = \frac{1}{\psi}$, recursive preferences reduce to constant-relative risk aversion (CRRA) preferences.

Following Albuquerque et al. (2015), I model demand-side shocks to asset prices by allowing stochastic changes in the agent’s time discount rate. The ratio $\delta \cdot (\lambda_{t+1}/\lambda_t)$ determines how agents trade off current period versus future period utility, which I assume is known at time $t$. I refer to $x_{\Lambda,t} \equiv \log(\lambda_{t+1}/\lambda_t)$ as a time preference shock, which evolves according to

$$x_{\Lambda,t+1} = \rho_{\Lambda} \cdot x_{\Lambda,t} + \sigma_{\Lambda} \cdot \varepsilon_{\Lambda,t+1}$$

where $\varepsilon_{\Lambda,t+1}$ is an i.i.d. standard-normal random variable.

Economic fundamentals, denoted by $z_{t+1}$, include consumption growth, dividends growth and inflation, i.e., $z_{t+1} \equiv (\Delta c_{t+1}, \Delta d_{t+1}, \pi_{t+1})'$. I specify that $E_t(z_{t+1})$, the conditional mean of economic fundamentals is linear in the state variable $x_{z,t} \equiv (x_{c,t}, x_{\pi,t})'$, such that $E_t[z_{t+1}] = h_x \cdot x_{z,t}$. The vector $x_{z,t+1}$ is modeled as a VAR(1) process,

$$\begin{bmatrix} x_{c,t+1} \\ x_{\pi,t+1} \end{bmatrix} = \begin{bmatrix} \rho_x & 0 \\ 0 & \rho_{\pi} \end{bmatrix} \begin{bmatrix} x_{c,t} \\ x_{\pi,t} \end{bmatrix} + \begin{bmatrix} 1 & \Phi_{x\pi} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_{x,t+1}^x \\ \eta_{t+1}^{\pi} \end{bmatrix}. \hspace{1cm} (4)$$

Note that $x_{c,t+1}$ is the long-run risk component which drives the conditional mean of consumption and dividend growth; while $x_{\pi,t}$ controls the conditional mean of inflation. Parameters $\rho_x$ and $\rho_{\pi}$ govern the persistence of $x_{c,t+1}$ and $x_{\pi,t+1}$, respectively.\footnote{The autoregressive coefficient matrix is assumed to be diagonal, so the persistent component of inflation $x_{\pi,t}$ does not directly drive the long-run risk factor $x_{c,t+1}$.} Contemporaneous shocks $\eta_{t+1}^x$ and $\eta_{t+1}^{\pi}$, are assumed to be i.i.d normally distributed. Parameter $\Phi_{x\pi}$
allows shocks to the persistent component of inflation \( \eta_{t+1}^\pi \) to affect the the long-run risk component \( x_{z,t+1} \). Given the definition of \( x_{z,t+1} \), the exogenous process for \( z_{t+1} \) is modeled as follows,

\[
\begin{align*}
\Delta c_{t+1} &= \mu_c + \pi_{ct} \cdot \varepsilon_{\Lambda t+1}^A + \Phi_{c\pi} \cdot \eta_{t+1}^\pi + \eta_{t+1}^c \\
\Delta d_{t+1} &= \mu_d + \phi_d \cdot \pi_{ct} + \pi_{dt} \cdot \varepsilon_{\Lambda t+1}^A + \Phi_{d\pi} \cdot \eta_{t+1}^\pi + \pi_{dc} \cdot \eta_{t+1}^c + \eta_{t+1}^d \\
\pi_{t+1} &= \mu_{\pi} + x_{\pi,t} + \eta_{t+1}^\pi
\end{align*}
\]

where \( \eta_{t+1}^c, \eta_{t+1}^d \) and \( \eta_{t+1}^\pi \) are i.i.d normal shocks to the transitory components of consumption growth, dividends growth and inflation, respectively. Preference shocks \( \varepsilon_{t+1}^\Lambda \) and shocks to the persistent component of inflation \( \eta_{t+1}^\pi \) also spill over to consumption and dividend growth.

There is considerable evidence in favor of stochastic volatilities in post-war macroeconomic and financial data.\(^2\) I assume that all supply-side shocks have stochastic volatilities, and model those shocks as

\[
\eta_{t+1}^j \equiv \sigma_{jt} \cdot \varepsilon_{t+1}^j, \quad \text{for } j \in \{x, \pi, c, d, \iota\}
\]

where each \( \varepsilon_{t+1}^j \) is an i.i.d. standard-normal random variable, and each \( \sigma_{jt+1}^2 \) follows an autoregressive gamma (ARG) Process, introduced by Gourieroux and Jasiak (2006). In general, a non-negative scalar time series \( h_{t+1} \) with an ARG \((\phi, \nu, c)\) representation evolves over time according to the following relationship,

\[
\begin{align*}
h_{t+1} &\sim \text{Gamma}(\nu + \varsigma_{t+1}, c) \\
\varsigma_{t+1} &\sim \text{Poisson}\left(\phi \cdot h_t / c\right),
\end{align*}
\]

where parameters \( \phi, \nu, \) and \( c \) correspond to the persistence, shape and scale of \( h_{t+1} \), respectively. Moreover, the conditional mean and conditional variance of \( h_{t+1} \) evolve according to,

\[
\begin{align*}
E_t[h_{t+1}] &= \nu c + \phi \cdot h_t \\
\text{Var}_t[h_{t+1}] &= \nu \cdot c^2 + 2 \cdot \phi c \cdot h_t
\end{align*}
\]

As shown in Le, Singleton, and Dai (2010), Hsu and Palomino (2015) and Creal and Wu (2016), this ARG stochastic volatility process has the desirable feature of both guaranteeing positive realizations of volatility states and making the solution tractable in this class of models.\(^3\) Appendix A contains more details about this ARG process.


\(^3\)Many previous works such as Bansal, Kiku, and Yaron (2012) and Albuquerque et al. (2015), specify
2.2 Solution for asset prices.

I define equity returns as the return to a claim on the dividend process,
\[ R_{d,t+1} \equiv \exp (r_{d,t+1}) = \frac{P_{d,t+1} + D_{t+1}}{P_{d,t}}, \]  
(6)

where \( P_{d,t} \) is the ex-dividend stock price. In order to solve for the agent’s stochastic discount factor, it is also useful to define the return to the endowment process as
\[ R_{c,t+1} \equiv \exp (r_{c,t+1}) = \frac{P_{c,t+1} + C_{t+1}}{P_{c,t}} \]  
(7)

where \( P_{c,t} \) denotes the price of an asset that pays a dividend equal to aggregate consumption. The representative agent’s utility maximization problem implies that the logarithm of the stochastic discount factor is given by
\[ m_{t+1} = \theta \log (\delta) + \theta \log \left( \frac{\lambda_{t+1}}{\lambda_t} \right) - \theta \cdot \Delta t_{t+1} + (\theta - 1) \cdot r_{c,t+1} \]  
(8)

and \( \theta \) is defined as
\[ \theta = \frac{1 - \gamma}{1 - 1/\psi}. \]

In the case of CRRA preferences, \( \gamma = 1/\psi \), and \( \theta \) is equal to one. In that case, the SDF is independent of \( r_{c,t+1} \).

I follow Campbell and Shiller (1988) and apply log-linear Taylor expansion to equations (6) and (7)
\[ r_{c,t+1} = \kappa_{c0} + \kappa_{c1} \cdot z_{c,t+1} - z_{c,t} + \Delta t_{t+1} \]  
(9)
\[ r_{d,t+1} = \kappa_{d0} + \kappa_{d1} \cdot z_{d,t+1} - z_{d,t} + \Delta d_{t+1} \]  
(10)

where the log price-consumption ratio and log price-dividend ratio are defined as \( z_{c,t+1} \equiv P_{c,t+1}/C_{t+1} \) and \( z_{d,t+1} \equiv P_{c,t+1}/C_{t+1} \). The linear coefficients \((\kappa_{c0}, \kappa_{c1})\) and \((\kappa_{d0}, \kappa_{d1})\) in equations (9) and (10), respectively, are functions of the unconditional means of \( z_{c,t} \) and \( z_{d,t} \), respectively.

The Euler equations associated with \( r_{c,t+1} \) and \( r_{d,t+1} \) can be written as
\[ 1 = E_t \left[ \exp (m_{t+1} + r_{c,t+1}) \right] \]  
(11)
\[ 1 = E_t \left[ \exp (m_{t+1} + r_{d,t+1}) \right] \]  
(12)

The stochastic volatility variable \( \sigma_t^2 \) as a Gaussian AR(1) process, such that
\[ \sigma_{t+1}^2 = (1 - \phi) \sigma^2 + \phi \cdot \sigma_t^2 + \sigma_\omega \cdot \omega_{t+1} \]

where \( \omega_{t+1} \) is an i.i.d standard normal random variable. This specification makes the solution tractable. However, \( \sigma_t^2 \) can possibly take on negative values, which is logically impossible.
I solve the model using the method of undetermined coefficients. First, I replace \( r_{c,t+1} \), \( r_{d,t+1} \) and \( m_{t+1} \) in equations (11) and (12) using expressions (9), (10) and (8). Then I guess and verify that the equilibrium solutions for \( z_{c,t+1} \) and \( z_{d,t+1} \) are affine functions of the state variable \( S_t \), such that

\[
\begin{align*}
    z_{c,t} &= A_{c0} + A'_{c} \cdot S_t \\
    z_{d,t} &= A_{d0} + A'_{d} \cdot S_t
\end{align*}
\]

where the economic state variable \( S_t \) consists of valuation risk component \( x_{\Lambda,t} \), long-run risk component \( x_{c,t} \), expected inflation \( x_{\pi,t} \), and as finally the five stochastic volatility variables,

\[
S_t = \begin{bmatrix}
    x_{\Lambda,t} & x_{c,t} & x_{\pi,t} \\
    \sigma^2_{x,t} & \sigma^2_{\pi,t} & \sigma^2_{c,t} & \sigma^2_{d,t} & \sigma^2_{\iota,t}
\end{bmatrix}.
\]

Solving this model also requires: (1) exploring the properties of log-normal distribution and the ARG process\(^4\) (2) numerically solving for the unconditional mean of \( z_{c,t} \) and \( z_{d,t} \). More details on the model solution are in appendix B.

Given the solution for \( m_{t+1} \), the short-term real interest rate \( r_{f,t+1} \) can be derived from the Euler equation, \( E_t [\exp (m_{t+1} + r_{f,t+1})] = 1 \) and is affine in \( S_t \),

\[
r_{f,t} = \Psi^{(1)}_0 + \Psi^{(1)}_S \cdot S_t
\]

The price of a nominal bond maturing in \( h \) periods, denoted as \( p_{t}^{(h)} \), can be derived recursively given

\[
p_{t}^{(h)} = \log E_t \left[ \exp \left( m_{t+1} \right) \right].
\]

Yields on these multi-period nominal bonds are then defined as \( y_{t}^{(h)} = -\frac{1}{n} \cdot p_{t}^{(h)} \) and are also affine in \( S_t \),

\[
y_{t}^{(h)} = \Psi^{(h)}_0 + \Psi^{(h)}_S \cdot S_t.
\]

\(^4\)By definition, if \( h_t \) is an ARG \((\phi, \nu, c)\) process, the Laplace transform of \( h_{t+1} \) conditional on \( h_t \) satisfies

\[
E_t [\exp (u \cdot h_{t+1})] = \exp \left[ -\nu \cdot \log (1 - u \cdot c) + \frac{u \cdot \phi}{1 - u \cdot c} \cdot h_t \right].
\]
2.3 Economic Intuitions

I define the vector of asset prices as

$$p_{t+1} = \left( z_{d,t}, r_{f,t}, \left\{ y_t^{(h)} \right\}_{h \in H} \right)'$$

Combining the pricing results derived above gives us

$$p_t = \Psi_0 + \Psi_{1,X} \cdot \chi_t + \Psi_{1,V} \cdot V_t.$$  \hspace{1cm} (13)

That is, all asset prices are affine in

$$S_t = \left[ x_{\Lambda,t}, x_{z,t}', V_t' \right]' .$$

$S_t$ are economic factors, where variable $x_{\Lambda,t}$ stems from the demand-side of the economy; in contrast, $x_{z,t}$ and $V_t$ are supply-side shocks. Asset prices are determined by both demand and supply, and coefficient matrices $\Psi_0$, $\Psi_{1,X}$ and $\Psi_{1,V}$ are non-linear functions of model parameters $\Theta$.

The expected return to an asset reflects the covariance between the asset’s payoff and the agent’s stochastic discount factor. Given the relatively smooth economic fundamentals, a highly volatile stochastic discount factor $m_{t+1}$ is needed to account for high equity returns. Under Epstein-Zin preferences, the agent is sensitive to shocks in the future and prefers early resolution of uncertainty. The agent fears both a long period of persistently low consumption growth (supply-side shocks) and persistently low consumption valuations (demand-side shocks). Both channels make the representative agent dislike holding stocks, which promise stochastic payoff for an infinite number of periods. In equilibrium this leads to high stock returns. In contrast, short-term real bonds are less exposed to these low-frequency shocks.

Although both demand and supply-side shocks can explain the equity premium, they generate opposing predictions for the shape of the real yield curve. The long-run risk model alone implies a downward-sloping real yield curve, whereas valuation risk implies an upward-sloping real yield curve. Consider an environment in which there are only long-run risk shocks; a, a risk-averse agent who prefers early resolution of uncertainty can buy long-term real bonds to hedge that risk. The risk of a persistent shock increases with horizon, making long-term bonds more attractive than short-term bonds and implying a downward-sloping real yield curve. When valuation risk is present, a risk-free bond no longer provides a guaranteed payoff in utility terms. This is because the marginal utility of consumption is not just a function of the consumption promised by the bond but also a function of the preference shock. Introducing valuation risk therefore reduces the insurance provided by real bonds, making them less valuable. Since uncertainty about preferences increases with maturity, this effect should be greater at longer maturities. The valuation risk model therefore implies an upward-sloping real yield curve. While TIPS data suggests the real yield curve is upward-sloping, lack of liquidity and a relatively short sample period make it difficult to draw definite conclusions. However, nominal yields are drawn from more liquid markets and can be used to derive the real yield curve once we understand the joint dynamics of inflation and consumption.
It is worth pointing out that both models fail to account for the near zero correlation between the price-dividend ratio \( z_{d,t} \) and the real interest rate \( r^f_t \). In the long-run risk model, when a positive shock increases the present value of dividends, stocks become more valuable compared to the risk-free bond. Therefore, the representative agent is willing to substitute stocks for bonds, resulting in a strongly positive correlation between the price-dividend ratio and the real interest rate. In contrast, when a preference shock makes the agent more patient in a valuation risk model, the agent wants to invest more in both stocks and bonds, generating a negative correlation. In my nested model, the correlation depends on the relative strength of these two opposing forces:

\[
\text{cov} \left( z_{d,t}, r^f_t \right) = -\frac{1}{1 - \kappa_d \cdot \rho_x} \cdot \sigma^2_x + \left( \phi_d - \frac{1}{\psi} \right) \cdot \frac{\sigma_d^2}{1 - \kappa_d \cdot \rho_x} \cdot \sigma^2_x + \text{other terms}. \tag{14}
\]

In equation (14), the loading on the variance of preference shock \( \sigma^2_{x,\lambda} \) is negative, while the loading on the unconditional variance of long-run risk shocks \( \sigma^2_x \) is positive. Therefore, my model allows these two different forces to balance each other out so as to better fit the data.

There are additional features of the data in which one model clearly outperforms the other. For instance, the long-run risk model is more capable of matching the stock return predictability by lagged price-dividend ratio. In contrast, the valuation risk model can explain the low correlation between asset returns and growth rate of economic fundamentals. While discussing specific moments can help to clarify the economic mechanism behind these two different types of risks, estimating the nested model allows for a more systematic comparison.

2.4 Inference

2.4.1 The data

The sample is 1952:06 to 2015:12. Consumption is defined as PCE on non-durable goods and services series. Since consumption series observed at a monthly frequency tend to suffer from measurement issues, as pointed out by Wilcox (1992), I use both monthly and quarterly consumption growth series for model inference. The SP500 stock price index and dividends data are from Robert Shiller’s website. I calculate inflation as the percentage changes in CPI, downloaded from FRED website. Short-term nominal interest rates with 1- and 3-month maturities, and Fama and Bliss (1987) zero-coupon yields on nominal bonds with maturities of 1 to 5 years are from CRSP. I obtain constant maturity interest rate data with maturities of 7 years, 10 years, 20 years and 30 years from Federal Reserve Tables H.15. The short-term real interest rate is constructed following procedures documented by Beeler and Campbell (2012).
2.4.2 Bayesian inference

I denote the history of a generic vector of variables $A_t$ by

$$A = [A_1, A_2, \ldots, A_T].$$

The vector $\Theta \equiv (\Theta_Y, \Theta_V)$ contains the structural parameters of the model, where $\Theta_V$ corresponds to the subset of model parameters that characterize the stochastic volatility processes and $\Theta_Y$ stacks all the other parameters. The observable vector $y_t$ consists of economic fundamentals $z_{t+1}$ and asset prices $p_{t+1}$. To facilitate inference, it is convenient to cast the model into the following non-linear state space representation:

$$y_{t+1} = C + H_X \cdot X_{t+1} + H_V \cdot V_{t+1} + \Gamma \cdot e_{t+1}$$  \hspace{1cm} (15)

$$X_{t+1} = F_X \cdot X_t + R(V_t) \cdot \varepsilon_{t+1}$$  \hspace{1cm} (16)

$$V_{t+1} \sim F_V(V_t; \Theta_V)$$  \hspace{1cm} (17)

where the elements of $e_{t+1}$ and $\varepsilon_{t+1}$ are i.i.d standard normal and uncorrelated with each other. The measurement equation (15) summarizes equations (3) and (13). The state variable $X_{t+1}$ consists of $\chi_{t+1}, \eta_{t+1}, \chi_t$ and their lagged values. The transition equations of $\chi_{t+1}$, defined in equations (3) and (4), are embedded in equation (16). Therefore, the innovation term in equation (16), $R(V_t) \cdot \varepsilon_{t+1}$, inherits the time-varying volatility of $\chi_{t+1}$. Finally, the elements of $V_{t+1}$ in (17) are uncorrelated with each other and are distributed according to an autoregressive gamma process. When estimating the model, I take into account of the deterministic changes in data availability and time aggregation, which requires a modification of the state space system.

All matrices in equations (15), (16) and (17) are functions of the parameters $\Theta$. The likelihood function of the non-linear state space system, $P(y | \Theta)$, is analytically intractable. Conditional on a history of $V$, however, equations (15) and (16) form a linear state-space system, so $P(y | \Theta, V)$, the model likelihood conditional on the history of $V$, is easily computed using the Kalman filter iteration.

I adopt a Bayesian approach to inference that combines relatively uninformative priors with likelihood information. I rely on Markov chain Monte Carlo (MCMC) methods to characterize the joint posterior distribution of the model parameters $(\Theta)$, the history of the state variables $(X)$ and the history of time-varying volatility states $(V)$;

$$P(\Theta, X, V | y).$$

This method efficiently deals with the high-dimensional parameter space and non-linearity of model state space by breaking the original estimation problem into several simpler blocks. Broadly, my MCMC algorithm is a three-step particle Gibbs sampling (PG) procedure. First, I use a standard random-walk Metropolis algorithm to sample structural parameters $\Theta$. Next, I sample $X$ by exploiting the conditional linear and Gaussian nature of the model.
given values of $V$. Finally, I draw a sample of time-varying volatilities $V$ from a collection of particles simulated from the particle filter.

To improve estimation efficiency, it is crucial to sample the history of stochastic volatilities $V$ within a single block. Justiniano and Primiceri (2008) achieve this by exploiting the block structure between $(\Theta_V, V)$ and the rest of their model. They then sample sequences of $V$ from an approximately linear Gaussian state space system using the method of Kim et al. (1998). However, that option is not available in my model, as the history of $V$ drives both the conditional mean and the conditional variance of $y$.

As a viable alternative, I use a particle Gibbs sampling procedure to draw $V$. The forward simulation algorithm of particle filters can approximate a sequence of continuous distribution of $\{P_t(V_t | \cdot)\}_{t=1}^T$ by a sequence of discrete distributions, each consisting of an $M$-tuple of support points and probability weights;

$$\left\{ V_t^{(m)}, \omega_t^{(m)} \right\}_{m=1}^M.$$  \hspace{1cm} (18)

The particle Gibbs sampler step, proposed by Andrieu, Doucet, and Holenstein (2010) and Whiteley (2010), takes the simulated distributions in (18) as candidate values, upon which the standard backward sampling method is applied to draw a path of $V$. Further details of the estimation are documented in Appendix C.

2.5 Estimation Results

The parameters are stacked into $\Theta \equiv (\Theta_Y, \Theta_V)$, where

$$\Theta_Y \equiv \left( \gamma, \psi, \delta, \rho, \sigma, \mu_c, \mu_d, \mu_\pi \right)^T \begin{pmatrix} \Phi_{cx} & \Phi_{dx} & \Phi_{dt} & \Phi_{xt} & \Phi_{xt} & \Phi_{xt} & \Phi_{xt} \\ \pi_{cx} & \pi_{dx} & \pi_{dt} & \pi_{xt} & \pi_{xt} & \pi_{xt} & \pi_{xt} \end{pmatrix} \begin{pmatrix} \lambda_c & \lambda_d & \lambda_p & \lambda_b \end{pmatrix}$$

$$\Theta_V \equiv \left( \phi_i, \sigma_i, s_i \right)_{i \in \{x, c, d, \pi\}}.$$

The observable vector for model estimation is as follows;

$$y_t = \left\{ \Delta c_t^q, \Delta c_t, \Delta d_t, \pi_t, z_{d,t}, r_{f,t}, r_{t}^S, \left\{ y_{t}^{(h)} \right\}_{h \in \{3, 12, 24, 36, 48, 60, 120, 240}} \right\}.$$  

Asset prices are assumed to be observed with i.i.d. normal measurement errors.

2.5.1 Priors

Prior distribution of $\Theta_Y$. I place uninformative priors on the agent’s preference parameter. The risk aversion parameter $\gamma$ is centered around 5 with a standard deviation of 2.5, and the elasticity of inter-temporal substitution parameter $\psi$ is centered around 1.5
with a standard deviation of 0.25. I calibrate the subjective discount rate $\delta$ to be 0.999. The unconditional variance of the time preference shock $x_{\Lambda,t}$, defined as $\sigma^2_{x_{\Lambda}} \equiv \frac{\sigma^2}{1-\rho^2}$, is centered around 0.002 so the corresponding credible region for steady-state real interest rate ranges from -3.5% to 5.9%. I place a relatively agnostic prior on the persistence of the long-run risk component $\rho_x$ and the valuation risk component $\rho_\lambda$ by specifying a beta distribution with mean of 0.9 and standard deviation of 0.1 on both parameters. I calibrate $\rho_\pi$ the autocorrelation of the persistent component of inflation, to be 0.975. The prior distributions for parameters $\Phi_{x\pi}$, $\Phi_{c\pi}$, and $\Phi_{d\pi}$, which govern inflation spillovers to consumption growth and dividend growth, are also dispersive and centered around zero. $\Phi_d$ is the loading of consumption long-run risk $x_{c,t}$ onto dividend growth $\Delta d_{t+1}$, and I assume it covers the interval between 1 to 8. I calibrate $\mu_c$, $\mu_d$ and $\mu_\pi$ to be the unconditional empirical means of consumption growth, dividend growth and inflation. Parameters $(\lambda_p)^2$ and $(\lambda_b)^2$ are defined as the percentage of variation in the price-dividend ratio and bond yields that can be explained by i.i.d. normal measurement errors. I specify $\lambda_p$ and $\lambda_b$ both follow a beta prior distribution with mean of 0.1 and standard deviation of 0.1.

**Prior distribution of $\Theta_V$.** To facilitate model estimation, I re-parametrize the stochastic volatility processes $\sigma^2_{i,t+1}$ by three parameters, $\phi_i$, $\bar{\sigma}_i$ and $s_i$. I calibrate $\phi_i$, the parameter that controls the autocorrelation of $\sigma^2_{i,t+1}$, to be 0.995. The unconditional mean of the stochastic volatility is $(\bar{\sigma}_i)^2 \equiv E(\sigma^2_{i,t+1})$, and I place uninformative priors on it. Specifically, $\eta_{t+1}$ is specified so that the unconditional variance of the persistent component of inflation, $x_{\pi,t}$ accounts for roughly 60 percent to 90 percent of variation in inflation, while the variance of the transitory component of inflation, $\eta_{t+1}$, accounts for the complementary 10 percent to 30 percent. The transitory shocks to consumption growth, $\eta_{c,t+1}$, account for about 50% to 90% of total variance in consumption, while the transitory shocks to dividend growth, $\eta_{d,t+1}$, account for 50% to 90% of total variation of dividend growth. Finally, the parameter $s_i$ is the ratio of the standard deviation to the mean of stochastic volatility states, such that $s_i \equiv \frac{\sigma(\sigma^2_{i,t})}{E(\sigma^2_{i,t})}$. I calibrate this ratio to “reasonable” values based on a first-stage parameter estimation of my model, where only data on economic fundamentals $z_{t+1}$ are used for inference. This procedure extracts information on $s_i$ from $z_t$ only. As pointed out by [Duffee et al. (2013)](https://dx.doi.org/10.1257/aer.103.3.830) and others, joint estimation of bond prices and economic fundamentals typically generates good fit on asset prices, but at the expense of poor fit of economic fundamentals. This problem manifests itself in the sense that convergence on $s_i$ is very hard to achieve, even when I attempt to estimate these $s_i$’s with very tight priors. Intuitively, the higher the volatility of $\sigma^2_{i,t}$, the better the model’s ability to track the volatility of asset prices. Although high $s_i$ will imply the model does not have good fit on $z_{t+1}$ as economic series tend to be smooth and stochastic volatilities do move but not that much. Directly calibrating these $s_i$’s based on economic fundamentals alleviates the problem and ensures that other parameters in $\Theta$ can be jointly estimated with both economic fundamentals and asset prices.

The first three columns of Table 1 report the priors. In addition to placing priors on the
parameters, I follow [Dew-Becker (2014)] and incorporate prior information on data moments. Specifically, I impose restrictions so my estimation routine favors parameters that imply a steady-state term spread of 2 percent and an equity premium of 5.5%.

2.5.2 Posterior distribution of parameters

Posterior distributions of model parameters $\Theta$ are summarized in the last three columns of Table 1. Risk aversion is estimated to be around 6.23, and the EIS is very tightly distributed around 1.37. These estimates are consistent with those found in the literature. Both the long-run and valuation risk shocks are very persistent, with $\rho_\Lambda = 0.9940$ and $\rho_x = 0.9887$. $\Phi_{xx}$ controls the spill-over from inflation shock $\eta_{t+1}^\pi$ to the long-run risk component, and it is estimated to be negative at -0.22. The 90% credible region for $\Phi_{xx}$ ranges from -0.307 to -0.088. The spillover coefficients from preference shocks to consumption and dividend growth $\pi_{c\lambda}$ and $\pi_{d\lambda}$ have opposite signs, and the inflation spillover parameters $\Phi_{c\pi}$ and $\Phi_{d\pi}$ are also of different signs. Therefore, a preference shock or an inflation shock that temporarily increases consumption growth will depress dividend growth.
Table 1: Prior densities and posterior estimates

<table>
<thead>
<tr>
<th>Density</th>
<th>Mean</th>
<th>Std</th>
<th>Median</th>
<th>5%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>G</td>
<td>5</td>
<td>2.5</td>
<td>6.2279</td>
<td>6.0381</td>
</tr>
<tr>
<td>$\psi$</td>
<td>G</td>
<td>1.5</td>
<td>0.25</td>
<td>1.3698</td>
<td>1.3698</td>
</tr>
<tr>
<td>Coefficients on Valuation Risks $x_{A,t+1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_\lambda$</td>
<td>B</td>
<td>0.9</td>
<td>0.1</td>
<td>0.9940</td>
<td>0.9961</td>
</tr>
<tr>
<td>$\sigma_{x\lambda}$</td>
<td>IG</td>
<td>0.0015</td>
<td>0.00075</td>
<td>0.00096765</td>
<td>0.00095</td>
</tr>
<tr>
<td>$\pi_{c\lambda}$</td>
<td>N</td>
<td>0</td>
<td>1</td>
<td>9.4199e-05</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\pi_{d\lambda}$</td>
<td>N</td>
<td>0</td>
<td>1</td>
<td>-0.0007</td>
<td>-0.0011</td>
</tr>
<tr>
<td>Coefficients on Long-run Risk $x_{c,t+1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>B</td>
<td>0.9</td>
<td>0.1</td>
<td>0.9887</td>
<td>0.9743</td>
</tr>
<tr>
<td>$\sigma_x$</td>
<td>IG</td>
<td>0.27</td>
<td>0.54</td>
<td>0.2282</td>
<td>0.2209</td>
</tr>
<tr>
<td>Coefficients on inflation process $x_{\pi,t+1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>IG</td>
<td>0.72</td>
<td>1.44</td>
<td>0.4906</td>
<td>0.4740</td>
</tr>
<tr>
<td>$\Phi_{xx}$</td>
<td>N</td>
<td>0</td>
<td>5</td>
<td>-0.2203</td>
<td>-0.307</td>
</tr>
<tr>
<td>$\Phi_{cp}$</td>
<td>N</td>
<td>0</td>
<td>5</td>
<td>1.0262</td>
<td>0.852</td>
</tr>
<tr>
<td>$\Phi_{dp}$</td>
<td>N</td>
<td>0</td>
<td>5</td>
<td>-0.0766</td>
<td>-0.1054</td>
</tr>
<tr>
<td>Parameters controlling Consumption-Dividend Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Phi_d$</td>
<td>G</td>
<td>5</td>
<td>2.5</td>
<td>3.3454</td>
<td>3.3437</td>
</tr>
<tr>
<td>$\pi_{dc}$</td>
<td>N</td>
<td>0.25</td>
<td>0.125</td>
<td>0.8275</td>
<td>0.7981</td>
</tr>
<tr>
<td>Unconditional Standard Deviation of Shocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>IG</td>
<td>2</td>
<td>1</td>
<td>1.9686</td>
<td>1.8984</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>IG</td>
<td>5</td>
<td>2.5</td>
<td>4.4781</td>
<td>4.2251</td>
</tr>
<tr>
<td>$\sigma_\iota$</td>
<td>IG</td>
<td>1.85</td>
<td>0.37</td>
<td>1.638</td>
<td>1.6159</td>
</tr>
<tr>
<td>Scaled Standard Deviation of Measurement Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_b$</td>
<td>B</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3292</td>
<td>0.3274</td>
</tr>
<tr>
<td>$\lambda_p$</td>
<td>B</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0281</td>
<td>0.0275</td>
</tr>
</tbody>
</table>

In column titled ‘Density’, N stands for Normal, B Beta, G Gamma and IG inverse Gamma distribution. etc. Calibrated coefficients are $\delta = 0.999$, $\phi_i = 0.995$, for $i \in \{x, \pi, c, d, \iota\}$. $s_x = 0.1$, $s_\pi = 0.25$, $s_c = 0.08$, $s_d = 0.1$, $s_\iota = 0.15$, $\mu_c = 3.02$, $\mu_c = 1.8$, $\mu_\pi = 3.45$, and $\sigma_c^\iota = 0.1$. 


2.5.3 Posterior distribution of unobserved state variables

The posterior distribution of the economic state variable $S_t$ can also be extracted readily using my MCMC procedure. Figure 1 plots the time series of the posterior distribution of $\chi$ with NBER recession dates, and Figure 2 plots the posterior distribution of $V$. The solid lines correspond to posterior medians, and the 90 percent credible region is shaded grey. All series in $\chi$ are tightly estimated. The valuation risk shock displays a high level of persistence, and it also has a high correlation of 0.91 with the real interest rate. The long-run risk shocks $x_{c,t}$ are strongly pro-cyclical, peaking around 2000 and plunging rapidly during the Great Recession. The persistent component of inflation follows the usual pattern, peaking in the early 1980s and trending downwards ever since the Great Moderation.

The stochastic volatility series $\{V_t\}$ are estimated with more uncertainty, as the credible regions are much wider compared to those of $\chi_t$. The volatility state $\sigma^2_{x,t}$ does not display much time variation. The volatility state $\sigma^2_{\pi,t}$ displays a pattern similar to the one estimated in Stock and Watson (2007), peaking in in the early 1980s during the Great Inflation, and trending downwards ever since. The volatility states of idiosyncratic component of dividend growth and inflation, $\sigma^2_{d,t}$ and $\sigma^2_{\iota,t}$, both ticked up during the Great Recession of 2007 to 2009.

Figure 1: Smoothed Estimates of $\chi_t$
The estimated model displays great flexibility in fitting actual asset price data, as is evident in Figure 3.

2.5.4 Asset pricing moments

I report the model’s implications for the shape of the nominal yield curve in Table 2. Column (1) reports the unconditional mean of nominal yields, with Newey-West standard errors reported in the parentheses. Evaluated at the posterior median of my parameter estimates $\Theta^*$, the economic model implies an upward sloping nominal yield curve. However, if I remove the valuation risk component, the economic model implies a downward sloping nominal yield curve. Beeler and Campbell (2012) pointed out the real yield curve implied by the long-run risk model tend to be downward-sloping. Bansal and Shaliastovich (2013) also confirm this feature for the long-run risk model. My results here are consistent with Creal and Wu (2016), who also documented a downward sloping nominal yield curve for a long-run risk model. Results in this table show that valuation-risk shocks play an important role in accounting for the upward sloping nature of the nominal yield curve.

To assess the relative role played by valuation risk and long-run risk in determining variation in asset prices, I calculate the following statistics...
Figure 3: Model Implied Asset Prices vs. Actual Data

nominal bond yields: alternative maturities

Table 2: Unconditional Mean of the Nominal Yield Curve

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Nested Model</td>
<td>Excluding Valuation Risk</td>
</tr>
<tr>
<td>$E[r^{S}_{t}]$</td>
<td>4.67(0.60)</td>
<td>4.65</td>
<td>3.88</td>
</tr>
<tr>
<td>$E[y^{S(12)}_{t}]$</td>
<td>5.06(0.61)</td>
<td>4.88</td>
<td>3.90</td>
</tr>
<tr>
<td>$E[y^{S(60)}_{t}]$</td>
<td>5.69(0.58)</td>
<td>5.62</td>
<td>3.73</td>
</tr>
<tr>
<td>$E[y^{S(240)}_{t}]$</td>
<td>6.08(0.51)</td>
<td>6.22</td>
<td>1.67</td>
</tr>
</tbody>
</table>
Table 3: Unconditional Variance Accounted for by Valuation and Long-run Risk

<table>
<thead>
<tr>
<th></th>
<th>Demand-side (VR)</th>
<th>Supply-side ()</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock returns</td>
<td>46.76%</td>
<td>47.19%</td>
</tr>
<tr>
<td>3-month bond yield</td>
<td>59.85%</td>
<td>10.85%</td>
</tr>
<tr>
<td>1-year bond yield</td>
<td>60.69%</td>
<td>9.11%</td>
</tr>
<tr>
<td>5-year bond yield</td>
<td>62.12%</td>
<td>6.01%</td>
</tr>
<tr>
<td>20-year bond yield</td>
<td>65.53%</td>
<td>0.95%</td>
</tr>
</tbody>
</table>

\[
\frac{\sigma_{x,j,t}}{\sigma_{y,t}} = 1 - \frac{V(p_t | x_j,t = 0)}{V(p_t)},
\]

for both stock returns and nominal bond yields, and report them in Table 3. Both long-run and valuation risk components account for around 45 percent of total variation in stock returns. However, valuation risk is very important in determining the variation in nominal bond yields. The relative fraction accounted for increases with maturity, from 60% for the 3-month bond to 65% for the 20-year bond. In contrast, the variation in nominal bond yields accounted for by the long-run risk component decreases with maturity, declining from 11% for 3-month yields to 1% percent for 20-year yields. Overall, while both long-run and valuation risks are important in determining stock price movements, the valuation risk component is crucial in accounting for variation in nominal bond yields, especially at longer maturities.

3 Comparing Economic Factors with ATSM Factors.

In the previous section, I have related asset prices \( p_{t+1} \) to economic state variables \( S_t \) using a consumption-based model. I have also argued it is important to include nominal yields in estimation in order to evaluate the relative importance of valuation risks versus long-run risk. In this section, I explore the connection between \( S_t \) and statistic factors \( F_t \) from a canonical affine term structure model (ATSM). ATSMs are highly effective in tracking bond prices with as few as three factors; the downside is that the resulting factors have no formal economic interpretation. Specifically, an ATSM starts with a direct assumption that short-term nominal interest rate is an affine function of the unobserved factor \( F_t \),

\[
\begin{align*}
  r_t^{S(1)} &= \delta_0 + \delta_t' \cdot F_t \\
  F_t &= \mu + \Phi \cdot F_{t-1} + \Sigma \cdot \epsilon_t,
\end{align*}
\]

\[\epsilon_t \sim \mathcal{N}(0, I),\]

\[\text{In a future draft, I will discuss other empirical implications of the nested model.}\]
where according to equation (20), \( F_t \) follows a VAR(1) process. Long-term nominal bond prices are related to short-term bond prices via a no-arbitrage condition,

\[
p_t^{(n+1)} = \log E_t \left[ \exp \left( m_{t+1}^s + p_{t+1}^{(n)} \right) \right]
\]

Equation (21) defines the logarithm of the stochastic discount factor \( m_{t+1}^s \), where risk-averse agents price in shocks via the term \( \lambda_t' \cdot \varepsilon_{t+1} \). Equation (22) specifies that the time-varying risk premium \( \lambda_t \) is an affine function of \( F_t \). The setup of an ATSM consists of equations (19), (20), (21), (22), and (23). Yields on an \( n \)-period nominal bond are also affine in state variable \( F_t \):

\[
y_t^{(n)} = A_n + B_n' \cdot F_t,
\]

where \( A_n = -\frac{1}{n} \cdot \overline{A}_n \), and \( B_n = -\frac{1}{n} \cdot \overline{B}_n \). Coefficients \( \overline{A}_{n+1} \) and \( \overline{B}_{n+1} \) are defined recursively as

\[
\overline{A}_{n+1} = -\delta_0 + \overline{A}_n + \overline{B}_n' (\mu - \Sigma \lambda_0) + \frac{1}{2} \overline{B}_n ' \Sigma \Sigma' \overline{B}_n
\]

\[
\overline{B}_{n+1} = -\delta_1' + \overline{B}_n' (\Phi - \Sigma \lambda_1),
\]

with \( \overline{A}_1 = -\delta_0 \) and \( \overline{B}_1 = -\delta_1 \). A detailed solution is provided in appendix E. The state variable \( F_t \) consists of three unobserved factors. Following Ang and Piazzesi (2003) I assume \( \mu = 0 \), \( \Sigma = I_3 \), and that \( \Phi \) is a lower diagonal matrix with \( \Phi_{11} \geq \Phi_{22} \geq \Phi_{33} > 0 \). The model parameters of interest are

\[
\Psi \equiv [\delta_0, \text{vec} (\delta_1), \text{vec} (\lambda_0), \text{vec} (\lambda_1), \text{vec} (\Phi), \text{vec} (\sigma_{me})].
\]

I estimate parameters of this ATSM by numerically maximizing \( P (y \mid \Psi) \), the likelihood of observing nominal bond yields data \( y_t \), where

\[
y_t \equiv \begin{bmatrix} y_t^{(1)} & y_t^{(3)} & y_t^{(12)} & y_t^{(24)} & y_t^{(36)} & y_t^{(48)} & y_t^{(60)} \end{bmatrix}
\]

To facilitate computation of \( P (\Psi \mid y) \), I recast the ATSM into the following state space representation,

\[
y_t = A + B \cdot F_t + M \cdot e_t \quad \quad e_t \sim \mathcal{N} (0, I)
\]

\[
F_t = \mu + \Phi \cdot F_{t-1} + \Sigma \cdot \varepsilon_t, \quad \quad \varepsilon_t \sim \mathcal{N} (0, I)
\]

See Hamilton and Wu (2012) for alternative specification of ATSMs.
where \( e_t \) is the vector of measurement errors. I assume that three nominal bond yields series \( (y_{t}^{(1)}, y_{t}^{(12)}, y_{t}^{(60)}) \) are observed without measurement error, and the other four bond yields are measured with i.i.d. normal errors. I calibrate \( \delta_0 \) to be the sample unconditional mean of short rate \( y_r^{(1)} \) and place restrictions in my estimation routine to search for a value of \( \Psi \) that matches the unconditional mean of 5-year rate \( y_r^{(60)} \). I label this particular ATSM as the benchmark model and report parameter estimates in Table 4. Note that consistent with Ang and Piazzesi (2003), \( F_{1,t} \) is highly persistent with \( \Phi_{22} = 0.990 \), \( F_{2,t} \) is also quite persistent with \( \Phi_{22} = 0.95 \), and the least persistent factor \( F_{3,t} \) is mean-reverting with \( \Phi_{33} = 0.68 \). These state variables \( (F_{1,t}, F_{2,t} \text{ and } F_{3,t}) \) can be readily interpreted as the “level”, “slope” and “curvature” factors, named after the different initial impacts they have on the nominal yield curve. Figure 4 plots the coefficient matrix \( B_n \) as a function of months to maturity \( n \). The coefficient on the level factor \( F_{1,t} \) is relatively flat, the coefficient on the slope factor \( F_{2,t} \) is upward sloping, and finally the coefficient on the curvature factor \( F_{3,t} \) is hump shaped. Therefore, on impact, \( F_{1,t} \) moves the entire yield curve about the same amount, \( F_{2,t} \) primarily affects the short end of the yield curve and \( F_{3,t} \) moves the short end and middle end of the yield curve in opposite directions.

### Table 4: ATSM Estimates: The Benchmark Model

<table>
<thead>
<tr>
<th>Short rate parameter ( \delta_1 \times 100 )</th>
<th>AR(1) coefficient matrix ( \Phi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{1,t} )</td>
<td>0.027</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>( F_{2,t} )</td>
<td>-0.020</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>( F_{3,t} )</td>
<td>0.030</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.021)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Price of risk ( \lambda_0 ) and ( \lambda_1 )</th>
<th>( \lambda_1 ) matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_{1,t} )</td>
<td>-0.141</td>
</tr>
<tr>
<td>(0.012)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>( F_{2,t} )</td>
<td>0.082</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>( F_{3,t} )</td>
<td>-0.176</td>
</tr>
<tr>
<td>(0.045)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

21
3.1 Correlations between $S_t$ and $F_t$

I document correlations between extracted economic factors and affine factors from the benchmark ATSM. The main message here is that preference shocks, defined as $-x_{\Lambda,t+1}$, are more strongly correlated with affine factors than the long-run risk factors, $x_{c,t+1}$. As reported in Table (5) below, the correlations between preference shocks and level, slope and curvature factors are 0.87, 0.39 and 0.61, respectively, while the correlations between long-run risk and affine factors are -0.10, -0.08 and -0.09. Note that $x_{\pi,t+1}$ captures agent’s inflation expectations, and therefore, consistent with the literature, it also has a high correlation (0.81) with the level factor $F_{1,t}$ of 0.81.

Table 5: Correlations between Economic Factors and Affine Factors.

<table>
<thead>
<tr>
<th>Estimated Economic Factors $S_t$</th>
<th>$-x_{\Lambda,t}$</th>
<th>$x_{c,t}$</th>
<th>$x_{\pi,t}$</th>
<th>$\sigma_{\pi,t}^2$</th>
<th>$\sigma_{\pi,t}^2$</th>
<th>$\sigma_{\pi,t}^2$</th>
<th>$\sigma_{d,t}^2$</th>
<th>$\sigma_{e,t}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{1,t}$ (Level)</td>
<td>0.87</td>
<td>-0.10</td>
<td>0.81</td>
<td>0.72</td>
<td>0.46</td>
<td>0.61</td>
<td>-0.74</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.02)</td>
<td>(0.004)</td>
<td>(0.12)</td>
<td>(0.17)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>$F_{2,t}$ (Slope)</td>
<td>0.39</td>
<td>-0.08</td>
<td>0.21</td>
<td>0.32</td>
<td>0.08</td>
<td>0.11</td>
<td>-0.35</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.02)</td>
<td>(0.007)</td>
<td>(0.14)</td>
<td>(0.15)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>$F_{3,t}$ (Curvature)</td>
<td>0.61</td>
<td>-0.09</td>
<td>0.61</td>
<td>0.50</td>
<td>0.30</td>
<td>0.30</td>
<td>-0.48</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.02)</td>
<td>(0.005)</td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.10)</td>
</tr>
</tbody>
</table>

Additionally, I investigate which elements in $S_t$ are crucial in driving moments in the affine factors $F_t$. According to the consumption-based model, nominal bond yields are affine
in economic factors:

\[ y_t^{(n)} = \Psi_0^{(n)} + \Psi_S^{(n)} \cdot S_t. \quad (26) \]

I construct an alternative dataset of bond yields \( \tilde{y}_t^{(n)} \) by plugging estimated economic factors \( \tilde{S}_t \) into equation (26), but with one of the factors in \( S_t \) set to zero. I then re-estimate the ATSM using the simulated nominal yields \( \{ \tilde{y}_t^{(n)} \}_{n \in N} \). Maximum likelihood estimation delivers parameter estimates \( \tilde{\Psi} \) and affine factor estimates \( \tilde{A}_n, \tilde{B}_n \), such that

\[ \tilde{y}_t^{(n)} = \tilde{A}_n + \tilde{B}_n \cdot \tilde{F}_t. \]

If \( \tilde{F}_t \) are highly correlated with the original affine factors \( F_t \), then the removed economic factor does not play a large role in accounting for the properties of the statistical affine factors. Therefore, the correlation between \( \tilde{F}_t \) and \( F_t \) is a measure of how intact the affine factor remains if a particular economic factor is excluded from its inference. The lower that correlation, the larger the role the excluded economic factor plays. Table 6 below reports the correlation between \( F_t \) and the alternative series of \( \tilde{F}_t \), constructed by sequentially setting each element of \( S_t \) to zero and applying the above mentioned procedure. In comparison with all other economic factors, the removal of the valuation risk component (setting \( x_{\Lambda,t+1} = 0 \)) yields the lowest correlation between implied affine factors and original affine factors. The correlation is 0.79 for level factors, 0.16 for slope factors and 0.60 for curvature factors. Consistent with prior research linking inflation measures to the level factor, the removal of the expected inflation factor (setting \( x_{\pi,t+1} = 0 \)) results in the second lowest correlation between the original and implied level factor. By comparison, the removal of the long-run risk factor (setting \( x_{c,t+1} = 0 \)) leaves the level factor more or less intact; the implied correlation are 0.93, 0.34 and 0.66, respectively. I take results in Table 5 and Table 6 as evidence that the valuation risk component is closely connected to statistical factors from affine term structure models.

Table 6: Correlations between Original Affine Factors \( F_t \) and Implied Affine Factors \( \tilde{F}_t \)

<table>
<thead>
<tr>
<th>Economic Factor Excluded (( S_j,t = 0 ))</th>
<th>( x_{\Lambda,t} )</th>
<th>( x_{c,t} )</th>
<th>( x_{\pi,t} )</th>
<th>( \sigma^2_{x,t} )</th>
<th>( \sigma^2_{\pi,t} )</th>
<th>( \sigma^2_{c,t} )</th>
<th>( \sigma^2_{d,t} )</th>
<th>( \sigma^2_{\iota,t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr (( F_1,t, \tilde{F}_1,t ))</td>
<td>0.79</td>
<td>0.93</td>
<td>0.90</td>
<td>0.94</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>corr (( F_2,t, \tilde{F}_1,t ))</td>
<td>0.16</td>
<td>0.34</td>
<td>0.34</td>
<td>0.38</td>
<td>0.34</td>
<td>0.33</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>corr (( F_3,t, \tilde{F}_3,t ))</td>
<td>0.60</td>
<td>0.66</td>
<td>0.61</td>
<td>0.66</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
</tr>
</tbody>
</table>
3.2 Affine Term Structure Models with Macro Variables

3.2.1 Estimated Coefficients $\Psi$

In this subsection, I follow Ang and Piazzesi (2003) closely and treat the estimated variables $\chi_t$ as observable factors and incorporate them into the standard ATSM. To start, I rewrite the short-rate equation (19) by including an additional economic observable factor $x_{j,t}$:

$$r^s_t = \delta_0 + \delta_{11} \cdot x_{j,t} + \delta_{12} \cdot F_t. \quad (27)$$

The state transition equation for this economy is therefore

$$\tilde{F}_t = \Phi \tilde{F}_{t-1} + \Sigma \varepsilon_t$$

where $x_{j,t}$ is modeled as an AR(1) process. I also retain the previous assumptions on $\Phi$ and $\Sigma$ and assume $\tilde{\Phi}$ to be block-diagonal. I do so in order to preserve the original three-factor interpretation of $F_t$ and to see how independent economic factors can change the interpretation of the original factors. The vector $\Sigma_j$ allows the shocks to observable economic factor $x_{j,t}$ to spill over to $F_t$. To complete the model, I re-define the risk premium equation as

$$\lambda_t = \begin{bmatrix} \lambda_0 \\ \lambda_{0,x} \end{bmatrix} + \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} F_t \\ x_{j,t} \end{bmatrix}. \quad (29)$$

Compared to the benchmark model, the modified price of risk vector $\lambda_t$ includes an additional element, $\lambda_{0,x} + \lambda_{1,x} \cdot x_{j,t}$. I also impose the assumption that the off-diagonal terms of $\lambda_t$ are zero to highlight the role $x_{j,t}$ plays as an independent factor. The modified ATSM therefore consists of equations (27), (28), (29), (21), and (22). I consider two variants of this modified ATSM. In the first model, labeled ATSM-VR, I include the valuation risk component and set $x_{j,t} = -x_{\Lambda,t}$. In the second model, labeled ATSM-LRR, I include the long-run risk component as an economic variable and set $x_{j,t} = x_{c,t}$. Further, I calibrate $\rho_j$ to be the value from the economic model, and I rescale the $x_{j,t}$ series so that $\sigma_j = 1$. In the case where $\Sigma_j = 0$, coefficients $\delta_{11}$ in the short-rate equation (27) can be consistently estimated by OLS. Table 7 reports estimated parameters of ATSM-VR in Panel A and estimated parameters of ATSM-LRR in Panel B.

AR(1) coefficient matrices $\Phi$

The properties of the parameter estimates are as follows. First, according to AR(1) coefficient matrices $\Phi$ of equation 28 reported in Table 7, the latent factor $F_{1,t}$ is highly persistent around 0.99, and $F_{2,t}$ is less so but still quite persistent (around 0.95), and the least persistent factor $F_{3,t}$ displays mean reversion. This pattern is true is both ATSM-VR and ATSM-LR, which is also consistent with estimates of $\Phi$ from the benchmark model. Second, both the valuation risk factor $x_{\Lambda,t}$ and the long-run risk factor $x_{c,t}$ enter the short-rate equation (27) in a statistically significant way. The coefficient estimates of $100 \times \delta_{11}$ are 0.0155 for ATSM-VR,
and -0.0071 for ATSM-LRR. The coefficient on $F_{1,t}$ in the short-rate equation is -0.005 and is statistically insignificant in the ATSM-VR model. In contrast, this coefficient is positive in both the original model (0.036) and the *ATSM-LR* model (0.027). I take this as evidence that including the valuation risk factor $x_{\Lambda,t+1}$ in ATSM-VR altered how the original “level factor” $F_{1,t}$ affects the short rate, while the inclusion of the long-run risk factor $x_{c,t+1}$ does not. Second, in the estimated equation for risk premium (29), the coefficient $\lambda_{0,x}$ is negative (-0.0337) for the ATSM-VR model and statistically zero for the ATSM-VR model. Ang and Piazzesi (2003) point out that negative values of $\tilde{\lambda}_0$ in equation (29) induce the nominal yield curve to be upward sloping on average. As in the consumption-based model, $x_{\Lambda,t}$ is crucial in generating an upward-sloping nominal yield curve here.

I plot in Figure 5, the factor loadings $B_n$ associated with ATSM-VR (Left Panel) and *ATSM-LR* (Right Panel), as a function of bond maturities $n$. I compare these loadings to those of the benchmark model in Figure 4. In contrast with that on $B_n$ from the benchmark model, the coefficient on the level factor in *ATSM-VR* is steeply upward-sloping, while the coefficient on valuation risk factor $-x_{\Lambda,t+1}$ resembles that of the “level factor” in Figure 4. Therefore, when valuation risk shocks are added to the ATSM as an independent variable, the modified ATSM-VR attributes a portion of the role played by the original “level factor” to the newly added variable, $-x_{c,t}$. In ATSM-LR, by comparison, coefficient loadings on “level”, “slope”, and “curvature” factors resemble those from the benchmark model.

**Figure 5: $B_n$ coefficients: ATSM-VR**
Table 7: Affine Term Structure Model with Macro Variables

### Panel A: ATSM-VR

<table>
<thead>
<tr>
<th>Short rate parameter $\delta_1 \times 100$</th>
<th>AR(1) coefficient matrix $\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{1,t}$ -0.005</td>
<td>0.9967 0 0</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.0035)</td>
</tr>
<tr>
<td>$F_{2,t}$ -0.039</td>
<td>-0.011 0.9318 0</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.007) (0.0788)</td>
</tr>
<tr>
<td>$F_{3,t}$ 0.019</td>
<td>-0.0017 -0.0034 0.670</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.020) (0.0211) (0.6677)</td>
</tr>
<tr>
<td>$-x_{\lambda,t}$ 0.0155</td>
<td></td>
</tr>
<tr>
<td>(0.0005)</td>
<td></td>
</tr>
</tbody>
</table>

Price of risk $\lambda_0$ and $\lambda_1$

| $F_{1,t}$ -0.0338                         | 0.0114 0.0124 -0.07211            |
| (0.0044)                                  | (0.0019) (0.0020) (0.020)         |
| $F_{2,t}$ 0.154                           | 0.03325 -0.016 -0.31242           |
| (0.0284)                                  | (0.0122) (0.0132) (0.0297)        |
| $F_{3,t}$ -0.0429                         | -0.0041 -0.0295 0.15643           |
| (0.0476)                                  | (0.0100) (0.0107) (0.0272)        |
| $-x_{\lambda,t}$ -0.03371                |                                 |
| (0.0097)                                  |                                 |

### Panel B: ATSM-LRR

<table>
<thead>
<tr>
<th>Short rate parameter $\delta_1 \times 100$</th>
<th>AR(1) coefficient matrix $\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{1,t}$ 0.036</td>
<td>0.98453 0 0</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.25818)</td>
</tr>
<tr>
<td>$F_{2,t}$ -0.017</td>
<td>0.022763 0.953743747 0</td>
</tr>
<tr>
<td>(0.009)</td>
<td>(0.0096905) (0.388141107)</td>
</tr>
<tr>
<td>$F_{3,t}$ 0.027</td>
<td>0.014662 0.0049117 0.757664752</td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.012729) (0.018274) (0.156953773)</td>
</tr>
<tr>
<td>$x_{c,t}$ -0.0071</td>
<td></td>
</tr>
<tr>
<td>(0.0026)</td>
<td></td>
</tr>
</tbody>
</table>

Price of risk $\lambda_0$ and $\lambda_1$

| $F_{1,t}$ -0.14                          | -0.024 -0.015418 0.11988          |
| (0.01)                                   | (0.007) (0.013235) (0.033879)     |
| $F_{2,t}$ 0.054                          | 0.038 -0.029291 -0.11561          |
| (0.01)                                   | (0.014) (0.016468) (0.020607)     |
| $F_{3,t}$ -0.11                          | -0.077 0.070372 0.2508            |
| (0.05)                                   | (0.023) (0.027295) (0.039219)     |
| $x_{c,t}$ 0.001                          | 26                               |
| (0.031)                                  |                                 |

$\lambda_0$ matrix

| $F_{1,t}$ -0.14                          | -0.024 -0.015418 0.11988          |
| (0.01)                                   | (0.007) (0.013235) (0.033879)     |
| $F_{2,t}$ 0.054                          | 0.038 -0.029291 -0.11561          |
| (0.01)                                   | (0.014) (0.016468) (0.020607)     |
| $F_{3,t}$ -0.11                          | -0.077 0.070372 0.2508            |
| (0.05)                                   | (0.023) (0.027295) (0.039219)     |
| $x_{c,t}$ 0.001                          | 26                               |
| (0.031)                                  |                                 |
3.2.2 Estimated Factors $\tilde{F}_t$

To formally investigate how the addition of the economic factors $x_{j,t}$ changes the latent factors ($F_{1,t}$, $F_{2,t}$ and $F_{3,t}$) of the benchmark model, I regress latent factors from the benchmark ATSM onto factors from the modified ATSMs. Table 8 reports the results of this exercise. Both the dependent variables and the regressors are normalized. Newey-West standard errors with 12 lags are reported in the parenthesis. The results of regression on ATSM-VR factors are reported in Panel A, while those of regression on ATSM-LR factors are reported in Panel B.

Table 8: Benchmark factors vs ATSM-VR and ATSM-LRR

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>Adj. $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark Factors</td>
<td>$x_{j,t}$</td>
<td>$F_{1,t}$</td>
<td>$F_{2,t}$</td>
<td>$F_{3,t}$</td>
<td>Adj. $R^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Regressions on ATSM-VR factors ($x_{j,t} = -x_{\Lambda,t}$)</td>
<td></td>
<td>0.484</td>
<td>0.497</td>
<td>-0.297</td>
<td>-0.030</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“level factor” $F_{1,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“slope factor” $F_{2,t}$</td>
<td>-0.531</td>
<td>1.340</td>
<td>0.754</td>
<td>0.050</td>
<td>0.9959</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“curvature factor” $F_{3,t}$</td>
<td>0.766</td>
<td>0.247</td>
<td>-0.743</td>
<td>0.840</td>
<td>0.9910</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Regressions on ATSM-LRR factors ($x_{j,t} = x_{c,t}$)</td>
<td></td>
<td>-0.07</td>
<td>0.976</td>
<td>0.038</td>
<td>-0.0023</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>“level factor” $F_{1,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0001)</td>
<td>(0.0002)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“slope factor” $F_{2,t}$</td>
<td>-0.024</td>
<td>-0.207</td>
<td>1.090</td>
<td>0.0134</td>
<td>0.9982</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.0025)</td>
<td>(0.0038)</td>
<td>(0.0039)</td>
<td>(0.0028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“curvature factor” $F_{3,t}$</td>
<td>-0.122</td>
<td>0.379</td>
<td>-0.0851</td>
<td>0.792</td>
<td>0.9959</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.0057)</td>
<td>(0.0059)</td>
<td>(0.0042)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For greater clarity, I will first explain results in Panel B. In this panel, I evaluate the effects of adding the long-run risk factor $x_{c,t}$ on the the latent factor. When I regress the level factor $F_{1,t}$ from the benchmark model onto ATSM-LR factors, the coefficient on $F_{1,t}$ is 0.976. Similarly, when the slope factor $F_{2,t}$ is regressed upon ATSM-LR factors, the coefficient on $F_{2,t}$ is 1.090. Both coefficients are close to unity. Thus, the level factor $F_{1,t}$ and the slope factor $F_{2,t}$ stay largely intact when the long-run risk factor $x_{c,t}$ is added to the ATSM model.
Panel A of Table 8 demonstrate the effect of adding the valuation risk factor $-x_{\Lambda,t}$ to the benchmark factor. In contrast with Panel B, when the level factor $F_{1,t}$ is regressed upon ATSM-VR factors, the coefficient on $F_{1,t}$ is only 0.484. When the slope factor $F_{2,t}$ is regressed upon ATSM-VR factors, the coefficient on $F_{2,t}$ is 0.754. Neither of these two coefficients are close to one. The intuition behind this change in the level factor when $-x_{\Lambda,t+1}$ is included can be shown from the following Taylor rule regression,

$$r_t^s = \beta_0 + \beta_1' \cdot X_t^o + \nu_t.$$  

When $X_t^o$ contains only the the long-run risk factor and the expected inflation factor, i.e. $X_t^o = [x_{c,t}, x_{\pi,t}]$. The autocorrelation of $\nu_t$ is 0.94 and the $R$ square is 0.71. When $-x_{\Lambda,t+1}$ is added to the regression, the autocorrelation of $\nu_t$ falls to 0.714 and $R$ square is increased to 0.96. Ang and Piazzesi (2003) mentioned that as long as the observable macro factor does not look like the interest rate itself, $\nu_t$ needs to be driven by some highly persistent level factor. When I use the identified preference shocks as an observable factor, given its resemblance to interest rate, it alters the role of the original level factor.

Although adding $x_{c,t}$ to ATSM does not alter the level and slope factors, the long-run risk factor $x_{c,t}$ does account for some movement in the level, slope, and curvature factors. The regression coefficients associated with $x_{c,t}$, as reported in column (2) of panel B, are all significant. The magnitudes of the effect that $x_{c,t}$ has on level, slope, and curvature factors are -0.07, -0.024, and -0.122, respectively. In comparison however, the effects of the valuation risk factor $x_{\Lambda,t}$ on the same factors are much greater, at 0.484, -0.531 and 0.766, respectively. This difference in magnitude suggests that valuation risk shocks affect ATSM to a much greater degree than long-run risk shocks.

In addition, I investigate how each estimated affine factor $F_{j,t}$ (or $\tilde{F}_{j,t}$) accounts for the total variance of nominal yields via the following variance decomposition experiment. I define $\text{frac}(n, j)$ as the fraction of total variance of $\hat{y}_t^{S(n)}$, which is unexplained if factor $F_{j,t}$ is set to zero:

$$\text{frac}(n, j) = 1 - \frac{V\left(\hat{y}_t^{S(n)} \mid F_{j,t} = 0\right)}{V\left(y_t^{S(n)}\right)}, \quad (30)$$

where $\hat{y}_t^{S(n)}$ is the ATSM-implied bond yields calculated based on equation (6). The advantage of this measure is that it is conditional on the state variable $F_t$ or $\tilde{F}_t$. Given that $F_{j,t}$’s are correlated with each other, the variance decomposition measure in equation (30) is not additive. In fact, $\text{frac}(n, j)$ can also take on negative values, in which case the ATSM-implied bond yields when $F_{j,t} = 0$ are more volatile than actual bond yields, i.e. $V\left(\hat{y}_t^{S(n)} \mid F_{j,t} = 0\right) > V\left(y_t^{S(n)}\right)$. Despite these potential issues, the magnitude of $\text{frac}(n, j)$ still sheds lights on the relative importance of each affine factors in determining nominal yields, conditional on the estimated state variable $F_t$. In Table 9 I report these $\text{frac}(n, j)$’s
for all $F_{j,t}$’s and selected maturities $(n)$. [1]

Table 9: Fraction of Nominal Yields Variance Explained by Affine Factors, frac(n, j)

<table>
<thead>
<tr>
<th>maturity in months</th>
<th>1</th>
<th>3</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Benchmark Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_{1,t}$</td>
<td>0.913</td>
<td>0.928</td>
<td>0.953</td>
<td>0.985</td>
<td>0.997</td>
<td>0.999</td>
<td>0.996</td>
</tr>
<tr>
<td>$F_{2,t}$</td>
<td>-0.182</td>
<td>-0.185</td>
<td>-0.149</td>
<td>-0.085</td>
<td>-0.027</td>
<td>0.020</td>
<td>0.057</td>
</tr>
<tr>
<td>$F_{3,t}$</td>
<td>0.287</td>
<td>0.133</td>
<td>-0.029</td>
<td>-0.047</td>
<td>-0.044</td>
<td>-0.038</td>
<td>-0.033</td>
</tr>
<tr>
<td>$-x_{\Lambda,t}$</td>
<td>0.578</td>
<td>0.592</td>
<td>0.603</td>
<td>0.591</td>
<td>0.563</td>
<td>0.530</td>
<td>0.496</td>
</tr>
</tbody>
</table>

| **Panel B: ATSM-VR** |       |       |       |       |       |       |       |
| $F_{1,t}$          | 0.202 | 0.212 | 0.349 | 0.506 | 0.615 | 0.691 | 0.747 |
| $F_{2,t}$          | 0.351 | 0.348 | 0.291 | 0.233 | 0.193 | 0.166 | 0.146 |
| $F_{3,t}$          | -0.164| -0.027| 0.081 | 0.079 | 0.067 | 0.058 | 0.050 |
| $-x_{\Lambda,t}$  | 0.578 | 0.592 | 0.603 | 0.591 | 0.563 | 0.530 | 0.496 |

| **Panel C: ATSM-LR** |       |       |       |       |       |       |       |
| $F_{1,t}$          | 0.918 | 0.924 | 0.953 | 0.985 | 0.993 | 0.991 | 0.985 |
| $F_{2,t}$          | -0.281| -0.272| -0.189| -0.081| -0.005| 0.072 | 0.123 |
| $F_{3,t}$          | 0.152 | 0.059 | -0.023| -0.030| -0.027| -0.023| -0.019|
| $x_{c,t}$          | 0.007 | 0.009 | 0.010 | 0.009 | 0.008 | 0.008 | 0.006 |

Panel A of Table 9 reports the unconditional variance decompositions for factors in the benchmark model. The “level factor” $F_{1,t}$ accounts for most of the variation of yields, and its importance increases with the maturity of yields. The statistics frac(n, j = 1) ranges from 0.913 for 1-month bond yields to 0.996 for 60-month yields. $F_{2,t}$ and $F_{3,t}$ jointly account for the remaining variation in bond yields, and the magnitude of frac(n, j = 2) and frac(n, j = 2) decreases with bond yield maturity. This pattern is broadly consistent with the fact that the “level factor” in ATSM is closely related to the first principle component of yield curve, which is known to account for over 90% of total variation in nominal yield curve.

Turning to Panel B of Table 9, I report the relative importance of each factor in ATSM-VR, which includes the preference shock series $-x_{\Lambda,t}$ an observable bond pricing factor. In this model, the “level factor” $F_{1,t}$ plays a much diminished role in explaining unconditional variances. The reading for frac(n, j = 1) is around 0.20 at the short end of the yield curve, climbing up to around 0.75 for a five-year nominal bond. These fractions are much lower than their counterparts in panel A. The inclusion of valuation risk shocks into standard ATSM also significantly increases the role played by the “slope factor” $F_{2,t}$. In ATSM-VR model, \footnote{In the appendix, I also report a proper unconditional variance decomposition that is conditional on the estimated model parameter $\Psi$ only. These results are qualitatively consistent with Table 9.}
accounts for 35% of 1-month bond yields, with its share dropping to 15% for 60-month bond yields. The declining influence of the “level factor” in ATSM-VR is primarily offset by the valuation risk shocks \(-x_{\Lambda,t}\). According to my model estimates, valuation risk shocks explain around 58% of total variance at the short end (1-month) of the nominal yield curve, and 50% at the long end (60-month) of the said curve. Therefore, when incorporated in a standard ATSM, valuation risk shocks account for a highly significant amount of variation of both the short end and the long end of the yield curve. The resemblance between the preference shock \(-x_{\Lambda,t+1}\) and the real interest rate plays a key role in ATSM-VR to accounting for variations at the long-end of the nominal yield curve. Similarly, early work by Evans and Marshall (2007) emphasizes the role of interest rate smoothing as an important channel for their model to link economic variables to the long end of yield curve.

Finally I investigate, in Panel C of Table 9, the relative importance of ATSM-LR factors. The first three rows of this panel are quantitatively similar to those in Panel A. Most of the variation in bond yields is accounted for by variation in the “level factor” \(F_{1,t}\). However, as reported in the fourth row, the long-run risk factor’s contribution to bond yields variations is below 1%. In a similar exercise, Ang and Piazzesi (2003) used both real economic activity and inflation as observable factors, modeled them as a VAR, and incorporated that process into a standard ATSM. Their approach allowed flexible correlation between real economic activities and inflation, and through the close connection between inflation and the “level factor” \(F_{1,t}\), attributed larger variation in yields to real economic activities. For this reason, numbers reported in the last role of Panel C potentially understate the importance of economic activity in determining variations of bond yields. However, the symmetry between ATSM-VR and ATSM-LR is preserved when I exclude measures of inflation from ATSM-LR.

In sum, when incorporated into the benchmark ATSM as an independent economic factor, the valuation risk shocks \(-x_{\Lambda,t}\): (1) significantly alter the original latent factors, especially the original “level factor” \(F_{1,t}\); and (2) account for a high fraction of unconditional variation in nominal yields, both at the short and long end of the nominal yield curve. In contrast, the effects are not as strong when the long-run risk shocks \(x_{c,t}\) are introduced as an independent economic factor into the benchmark ATSM.

4 Conclusion

This paper presents a consumption-based asset pricing model nesting both long-run and valuation risks. According to my model estimates, both risks are important in accounting for the behavior of stock returns and nominal bond yields. Valuation risk, modeled as persistent shocks to agents’ discount rates, plays a key role in accounting for the salient properties of the nominal yield curve. I also show that valuation risk shocks implied by this consumption-based model have a tight connection with statistical factors from a standard affine term structure model (ATSM). Although my economic model allows for shocks with stochastic volatility, the ATSM I study does not. A rich class of ATSM models allows for
heteroscedasticity of shocks (see Duffee (2002) and Dai and Singleton (2000), etc.). Treating stochastic volatility symmetrically in my economic model and the ATSMs is a natural next step.
References


33


Appendix

A Autoregressive Gamma Process

By definition, if $h_{t+1}$ follows an $ARG(\phi, \nu, c)$ process then

$$h_{t+1} \sim \text{Gamma}(\nu + \varsigma_{t+1}, c)$$

$$\varsigma_{t+1} \sim \text{Poisson} \left( \phi \cdot \frac{h_t}{c} \right).$$

This process guarantees the positive realization of volatilities, and has the following property where

$$E_t[h_{t+1}] = \nu c + \phi \cdot h_t$$

$$\text{Var}_t[h_{t+1}] = \nu \cdot c^2 + 2 \cdot \phi c \cdot h_t$$

The conditional density of this process is defined as

$$p(h_{t+1} \mid h_t) = h_t^{\nu-1} \cdot \frac{1}{c^{\nu}} \cdot \exp \left( -\frac{h_{t+1} + \phi \cdot h_t}{c} \right) \sum_{k=0}^{\infty} \frac{1}{\Gamma(\nu + k) \cdot c^k} \cdot h_{t+1}^k \cdot \left( \frac{\phi h_t}{c} \right)^k.$$ (31)

The expression for $p(h_{t+1} \mid h_t)$ involves summation of probability masses on the support of all positive integers. When computing this conditional density, I use the algorithm developed in [Le et al.] (2010) that provides an accurate approximation. When simulating realizations of $h_{t+1}$ given values of $h_t$, I take advantage of the fact that Poisson distribution with large arrival rate $\lambda$ converges to normal distribution with mean $\lambda$ and variance $\lambda$. The ARG process has the following property

$$E_t[\exp(u \cdot h_{t+1})] = \exp \left[ -\nu \cdot \log (1 - u \cdot c) + \frac{u \cdot \phi}{1 - u \cdot c} \cdot h_t \right].$$

which makes solution to asset prices affine in $h_t$. Unconditionally, $h_t$ follows a gamma distribution, with shape parameter $\nu$ and scale parameter $\frac{c}{1 - \phi}$. Therefore,

$$\mu_h \equiv E(h_t) = \nu \cdot \left( \frac{c}{1 - \phi} \right),$$

$$\sigma_h \equiv \sigma(h_t) = \sqrt{\nu} \cdot \left( \frac{c}{1 - \phi} \right).$$

When estimating the model, I re-parametrize the $ARG(\phi, \nu, c)$ process with parameters $\phi$, $\mu_h$, and $s_h \equiv \frac{\sigma_h}{\mu_h}$.  

35
B Solving the Consumption-based Asset Pricing Model.

I solve the Consumption-based Asset Pricing Model using the standard Campbell and Shiller (1988) log-linearization technique. In this framework, I define the gross return to consumption process and dividend process as,

\[ R_{c,t+1} = \frac{P_{c,t+1} + C_{t+1}}{P_{c,t}}, \quad R_{d,t+1} = \frac{P_{d,t+1} + D_{t+1}}{P_{d,t}} \]

where

\[ r_{c,t+1} = \log \left( R_{c,t+1} \right), \quad r_{d,t+1} = \log \left( R_{d,t+1} \right) \]
\[ z_{ct} = \log \left( \frac{P_{c,t}}{C_t} \right), \quad z_{dt} = \log \left( \frac{P_{d,t}}{D_t} \right) \]
\[ \Delta c_{t+1} = \log \left( \frac{C_{t+1}}{C_t} \right), \quad \Delta d_{t+1} = \log \left( \frac{D_{t+1}}{D_t} \right) \]

Log-linear Taylor expansion yields

\[ r_{c,t+1} = \kappa_{c0} + \kappa_{c1} \cdot z_{c,t+1} - z_{c,t} + \Delta c_{t+1} \]
\[ r_{d,t+1} = \kappa_{d0} + \kappa_{d1} \cdot z_{d,t+1} - z_{d,t} + \Delta d_{t+1}, \]

and the linear coefficients above are given by

\[ \kappa_{c1} = \frac{\exp (z_c)}{1 + \exp (z_c)}, \quad \kappa_{d1} = \frac{\exp (z_d)}{1 + \exp (z_d)} \]

and

\[ \kappa_{c0} = \log [1 + \exp (z_c)] - \kappa_{c1} \cdot z_c \]
\[ \kappa_{d0} = \log [1 + \exp (z_d)] - \kappa_{d1} \cdot z_d \]

where \( z_c \) and \( z_d \) are unconditional mean of \( z_{c,t} \) and \( z_{d,t} \).

This model solution has been discussed extensively by many previous authors, so I will be very brief here. Agent’s SDF is

\[ m_{t+1} = \theta \log (\delta) + \theta \cdot \log \left( \frac{\lambda_{t+1}}{\lambda_t} \right) - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \]

where

\[ r_{c,t+1} = \kappa_0 + \kappa_1 \cdot z_{c,t+1} - z_{c,t} + \Delta c_{t+1} \]

Transition equations for elements of \( \chi_{t+1} \) are

\[ x_{\Lambda,t+1} = \rho_\Lambda \cdot x_{\Lambda,t} + \sigma_{\Lambda} \cdot \xi_{t+1} \]
\[ x_{c,t+1} = \rho_x \cdot x_{c,t} + \sigma_{x,t} \cdot \xi_{t+1} + \Phi_{x\pi} \cdot \sigma_{\pi t} \cdot \xi_{t+1} \]
\[ x_{\pi,t+1} = \rho_\pi \cdot x_{\pi,t} + \sigma_{\pi t} \cdot \xi_{t+1} \]
and equations for the economic fundamentals are
\[
\begin{align*}
\Delta c_{t+1} &= \mu_c + x_{c,t} + \pi_{c,t} \cdot \varepsilon_{t+1}^c + \Phi_{cm} \cdot \sigma_{\pi,t} \cdot \varepsilon_{t+1}^\pi + \sigma_{c,t} \cdot \varepsilon_{t+1}^c \\
\Delta d_{t+1} &= \mu_d + \phi_d \cdot x_{c,t} + \pi_{d,t} \cdot \varepsilon_{t+1}^d + \Phi_{dm} \cdot \sigma_{\pi,t} \cdot \varepsilon_{t+1}^\pi + \pi_{dc} \cdot \sigma_{c,t} \cdot \varepsilon_{t+1}^c + \sigma_{dt} \cdot \varepsilon_{t+1}^d \\
\pi_{t+1} &= \mu_\pi + x_{\pi,t} + \sigma_{\pi,t} \cdot \varepsilon_{t+1}^\pi
\end{align*}
\]

B.1 Price Consumption Ratio

Conjecture that
\[
\begin{align*}
z_{c,t} &= A_0 + A_1 \cdot x_{\Lambda,t} + A_2 \cdot x_{c,t} + A_3 \cdot x_{\pi,t} + A_4 \cdot \sigma_{x,t}^2 + A_5 \cdot \sigma_{\pi,t}^2 + A_6 \cdot \sigma_{c,t}^2 + A_7 \sigma_{d,t}^2 + A_8 \sigma_{\iota,t}^2
\end{align*}
\]

The method of undetermined coefficients yields that
\[
\begin{align*}
A_1 &= \frac{1}{1 - \kappa_1 \rho_x}, & A_2 &= \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho_x}, & A_3 &= 0, & A_7 &= 0, & A_8 &= 0 \\
A_4 &= \frac{[1 + (\kappa_1 c_x) \cdot C_x - \kappa_1 \phi_x] - \sqrt{\Delta_x}}{2 (\kappa_1 c_x) \cdot \theta} \\
A_5 &= \frac{[1 + (\kappa_1 c_\pi) \cdot C_\pi - \kappa_1 \phi_\pi] - \sqrt{\Delta_\pi}}{2 (\kappa_1 c_\pi) \cdot \theta} \\
A_6 &= \frac{[1 + (\kappa_1 c_\iota) \cdot C_\iota - \kappa_1 \phi_\iota] - \sqrt{\Delta_\iota}}{2 (\kappa_1 c_\iota) \cdot \theta}
\end{align*}
\]

where
\[
\Delta_j = [1 + (\kappa_1 c_j) \cdot C_j - \kappa_1 \phi_j]^2 - 4 \cdot (\kappa_1 c_j) \theta \cdot C_j \quad j \in \{x, \pi, \iota\}
\]

and
\[
A_0 = \frac{1}{1 - \kappa_1} \left( \begin{array}{c}
\left[ \log (\delta) + \kappa_0 + \left( 1 - \frac{1}{\psi} \right) \mu_c \right] \\
+ \frac{1}{2} \left( \kappa_1 A_1 \sigma_\Lambda + \left( 1 - \frac{1}{\psi} \right) \pi_{c\Lambda} \right)^2 \\
- \nu_x \cdot \left( \frac{1}{2} \right) \cdot \log (1 - (\theta \kappa_1 A_4) \cdot c_x) \\
- \nu_\pi \cdot \left( \frac{3}{2} \right) \cdot \log (1 - (\theta \kappa_1 A_5) \cdot c_\pi) \\
- \nu_\iota \cdot \left( \frac{3}{2} \right) \cdot \log (1 - (\theta \kappa_1 A_6) \cdot c_\iota)
\end{array} \right)
\]

B.2 Price Dividend Ratio

Similarly, I guess and verify the solution to price dividend ratio is affine in state variables \( S_t \), where
\[
\begin{align*}
z_{d,t} &= A_{d0} + A_{d1} \cdot x_{\Lambda,t} + A_{d2} \cdot x_{c,t} + A_{d3} \cdot x_{\pi,t} + A_{d4} \cdot \sigma_{x,t}^2 + A_{d5} \cdot \sigma_{\pi,t}^2 + A_{d6} \cdot \sigma_{c,t}^2 + A_{d7} \cdot \sigma_{d,t}^2 + A_{d8} \cdot \sigma_{\iota,t}^2
\end{align*}
\]
\[ A_{d1} = \frac{1}{1 - \rho \kappa d1}, \quad A_{d2} = \frac{\phi d - \frac{1}{\phi}}{1 - \kappa d1 \rho}, \quad A_{d3} = 0, \quad A_{d8} = 0 \]

and \( A_{d4}, A_{d5}, A_{d6} \) and \( A_{d7} \) are solutions to the quadratic equation,

\[ A_j \cdot x^2 + B_j \cdot x + C_j = 0 \]

for \( j \in \{x, \pi, c, d\} \)

\[ A_j = c_j \cdot \kappa d1 \]
\[ B_j = -[(G_j - (\theta - 1)A_j) \cdot c_j \cdot \kappa d1 + (1 - c_j \cdot (\theta - 1) \kappa A_j) - \phi_j \kappa d1] \]
\[ C_j = [(G_j - (\theta - 1)A_j)(1 - c_j \cdot (\theta - 1) \kappa A_j) + \phi_j (\theta - 1) \kappa A_j] \]

where

\[ G_4 \equiv G_x = \frac{1}{2} [(\theta - 1) \kappa A_2 + \kappa d1 A_{d2}]^2 \]
\[ G_5 \equiv G_\pi = \frac{1}{2} [\begin{array}{c}
-\gamma \Phi c_\pi + \Phi d_\pi \\
+ [(\theta - 1) \kappa A_2 + \kappa d1 A_{d2}] \Phi x_\pi \\
+ [(\theta - 1) \kappa A_3 + \kappa d1 A_{d3}] \end{array}]^2 \]
\[ G_6 \equiv G_c = \frac{1}{2} (\pi_{dc} - \gamma)^2 \]
\[ G_7 \equiv G_d = \frac{1}{2} \]

and finally,

\[ (1 - \kappa d1) A_{d0} = \frac{1}{2} [\begin{array}{c}
\theta \log (\delta) + (\theta - 1) \kappa A_0 + (\theta - 1) (\kappa c_1 - 1) A_0 \\
-\gamma \mu c + \mu d \\
-\nu_c \cdot \log (1 - u_{d4} \cdot c_x) - \nu_\pi \cdot \log (1 - u_{d5} \cdot c) \\
-\nu_c \cdot \log (1 - u_{d6} \cdot c_x) - \nu_d \cdot \log (1 - u_{d7} \cdot c_d) \\
+ \frac{1}{2} (-\gamma \pi c_\lambda + \pi d_\lambda + [(\theta - 1) \kappa A_1 + \kappa d1 A_{d1}] \sigma A \end{array}]^2 \]

where

\[ u_{d,j} = (\theta - 1) \kappa A_j + \kappa d1 A_{dj}, \quad j = 4, 5, 6, 7 \]

### B.3 Real bond yields.

#### B.3.1 Real short-term interest rate.

By definition, the short-term real interest rate satisfies,

\[ r_{t}^{(1)} = -p_{t}^{(1)} = - \log E_t [\exp (m_{t+1})] \]
I guess and verify the solution to $p_t^{(1)}$ is of the form

$$p_t^{(1)} = p^{(1)} + B_1^{(1)} \cdot x, t + B_2^{(1)} \cdot x_c, t + B_3^{(1)} \cdot x_\pi, t + B_4^{(1)} \cdot \sigma_{x, t}^2 + B_5^{(1)} \cdot \sigma_{\pi, t}^2 + B_6^{(1)} \cdot \sigma_{c, t}^2$$

\[
B_1^{(1)} = 1 \\
B_2^{(1)} = -\frac{1}{\psi} \\
B_3^{(1)} = 0 \\
B_4^{(1)} = \frac{1}{2} [(\theta - 1) \cdot \kappa_1 \cdot A_2] + \frac{((\theta - 1) \cdot \kappa_1 \cdot A_4) \cdot \phi_x}{1 - ((\theta - 1) \cdot \kappa_1 \cdot A_4) \cdot c_x} - (\theta - 1) \cdot A_4 \\
B_5^{(1)} = \frac{1}{2} [-\gamma \Phi_c + (\theta - 1) \cdot \kappa_1 \cdot A_2 \Phi_x + (\theta - 1) \cdot \kappa_1 \cdot A_3] + \frac{((\theta - 1) \cdot \kappa_1 \cdot A_5) \cdot \phi_\pi}{1 - ((\theta - 1) \cdot \kappa_1 \cdot A_5) \cdot c_\pi} - (\theta - 1) \cdot A_5 \\
B_6^{(1)} = \frac{1}{2} \gamma^2 + \frac{((\theta - 1) \cdot \kappa_1 \cdot A_6) \cdot \phi_c}{1 - ((\theta - 1) \cdot \kappa_1 \cdot A_6) \cdot c_c} - (\theta - 1) \cdot A_6 \\
B_7^{(1)} = 0 \\
B_8^{(1)} = 0
\]

\[
p^{(1)} = \begin{cases} \\
\theta \log (\delta) + (\theta - 1) \kappa_0 - \gamma \mu_c + (\theta - 1) \cdot A_0 (\kappa_1 - 1) \\
+ \frac{1}{2} (\gamma \pi_{\lambda} + (\theta - 1) \cdot \kappa_1 \cdot A_1 \cdot \sigma_\lambda^2) \\
- \nu_x \cdot \log (1 - ((\theta - 1) \cdot \kappa_1 \cdot A_4) \cdot c_x) \\
- \nu_\pi \cdot \log (1 - ((\theta - 1) \cdot \kappa_1 \cdot A_5) \cdot c_\pi) \\
- \nu_c \cdot \log (1 - ((\theta - 1) \cdot \kappa_1 \cdot A_6) \cdot c_c) 
\end{cases}
\]

\[B.3.2 \quad \text{Yields on Long-term Bonds}\]

By definition,

$$p_t^{(n)} = E_t \left[ \exp \left( p_{t+1}^{(n-1)} + m_{t+1} \right) \right].$$

I guess and verify that,

$$p_t^{(n)} = p^{(n)} + B_1^{(n)} \cdot x, t + B_2^{(n)} \cdot x_c, t + B_3^{(n)} \cdot x_\pi, t + B_4^{(n)} \cdot \sigma_{x, t}^2 + B_5^{(n)} \cdot \sigma_{\pi, t}^2 + B_6^{(n)} \cdot \sigma_{c, t}^2.$$

Denote that

$$u_j^{(n)} = (\theta - 1) \cdot \kappa_1 \cdot A_j + B_j^{(n-1)} \quad j = 4, 5, 6$$
\[ p^{(n)} = p^{(n-1)} + \theta \log (\delta) + (\theta - 1) \kappa_0 - \gamma \mu_c + (\theta - 1) \cdot A_0 (\kappa_1 - 1) + \frac{1}{2} \left( -\gamma \pi + (\theta - 1) \cdot \kappa_1 \cdot A_1 \cdot \sigma + B_1^{(n-1)} \cdot \sigma \right)^2 \]
\[ -\nu_x \cdot \log \left( 1 - u_4^{(n)} \cdot c_x \right) - \nu_\pi \cdot \log \left( 1 - u_5^{(n)} \cdot c_\pi \right) - \nu_c \cdot \log \left( 1 - u_6^{(n)} \cdot c_c \right) \]
\[ B_1^{(n)} = \frac{1 - \rho^n_\Lambda}{1 - \rho_\Lambda} \]
\[ B_2^{(n)} = -\frac{1}{\psi} \left( \frac{1 - \rho_x^n}{1 - \rho_x} \right) \]
\[ B_3^{(n)} = 0 \]
\[ B_4^{(n)} = + \left[ \frac{1}{2} \left( (\theta - 1) \cdot \kappa_1 \cdot A_2 + B_2^{(n-1)} \right)^2 + \frac{u_4^{(n)} \cdot \phi_x}{1 - u_4^{(n)} \cdot c_x} - (\theta - 1) \cdot A_4 \right] \]
\[ B_5^{(n)} = + \left[ \frac{1}{2} \left( -\gamma \phi_x \Phi_x \pi + (\theta - 1) \cdot \kappa_1 \cdot A_2 \cdot \Phi_x \pi + B_2^{(n-1)} \cdot \Phi_x \pi \right)^2 + \frac{u_5^{(n)} \cdot \phi_\pi}{1 - u_5^{(n)} \cdot c_\pi} - (\theta - 1) \cdot A_5 \right] \]
\[ B_6^{(n)} = + \left[ \frac{1}{2} \gamma^2 - (\theta - 1) \cdot A_6 + \frac{u_6^{(n)} \cdot \phi_c}{1 - u_6^{(n)} \cdot c_c} \right] \]

**B.4 Nominal Bond Yields.**

**B.4.1 Nominal short-term interest rate**

By definition,
\[ p_t^{(1)} = \log E_t \left[ \exp \left( m_{t+1} - \pi_{t+1} \right) \right] \]
\[ p_t^{(1)} = p^{(1)} + B_t^{(1)} \cdot S_t \]

where
\[ p^{(1)} = p^{(1)} - \mu, \quad B_1^{(1)} = 1, \quad B_2^{(1)} = -\frac{1}{\psi}, \quad B_3^{(1)} = -1 \]
\[ B_4^{(1)} = B_4^{(1)}, \quad B_5^{(1)} = B_5^{(1)}, \quad B_6^{(1)} = B_6^{(1)}, \quad B_7^{(1)} = B_7^{(1)}, \quad B_8^{(1)} = B_8^{(1)} + \frac{1}{2} \]

Thus
\[ p_t^{(1)} = p_t^{(1)} - \mu - x_{\pi,t} + \frac{1}{2} \sigma_{\xi,t}^2 \]

**B.4.2 Yields on Long-term Nominal Bonds**

Denote
\[ u_i^{(n)} = (\theta - 1) \cdot \kappa_1 \cdot A_i + B_i^{(n-1)} \], \( i = 4, 5, 6, 7, 8 \)
Derivations yield that
\[ p^{S(n)} = p^{S(n-1)} + \theta \log(\delta) + (\theta - 1) \kappa_0 - \gamma \mu_c + (\theta - 1) \cdot A_0 (\kappa_1 - 1) \\
+ \frac{1}{2} \left( -\gamma \pi_c \lambda + (\theta - 1) \cdot \kappa_1 \cdot A_1 \cdot \sigma_\Lambda + B_1^{(n-1)} \cdot \sigma_\Lambda \right)^2 - \mu_\pi \\
- \nu_x \cdot \log \left( 1 - u_4^{S(n)} \cdot c_x \right) - \nu_\pi \cdot \log \left( 1 - u_5^{S(n)} \cdot c_\pi \right) \\
- \nu_c \cdot \log \left( 1 - u_6^{S(n)} \cdot c_c \right) - \nu_d \cdot \log \left( 1 - u_7^{S(n)} \cdot c_d \right) - \nu_\lambda \cdot \log \left( 1 - u_8^{S(n)} \cdot c_\lambda \right) \]
and that
\[ B_1^{S(n)} = \frac{1 - \rho_\Lambda^n}{1 - \rho_\Lambda} \]
\[ B_2^{S(n)} = - \frac{1}{\psi} \left( \frac{1 - \rho_\pi^n}{1 - \rho_\pi} \right) \]
\[ B_3^{S(n)} = - \left( \frac{1 - \rho_\pi^n}{1 - \rho_\pi} \right) \]
\[ B_4^{S(n)} = + \left[ \frac{1}{2} \left( (\theta - 1) \cdot \kappa_1 \cdot A_2 + B_2^{S(n-1)} \right)^2 + \frac{u_4^{S(n)} \cdot \phi_x}{1 - u_4^{S(n)} \cdot c_x} - (\theta - 1) \cdot A_4 \right] \]
\[ B_5^{S(n)} = + \left[ \frac{1}{2} \left( + (\theta - 1) \cdot \kappa_1 \cdot A_2 \cdot \Phi_{x\pi} + B_2^{S(n-1)} \cdot \Phi_{x\pi} \right)^2 + \frac{u_5^{S(n)} \cdot \phi_\pi}{1 - u_5^{S(n)} \cdot c_\pi} - (\theta - 1) \cdot A_5 \right] \]
\[ B_6^{S(n)} = + \left[ \frac{1}{2} \gamma + \frac{u_6^{S(n)} \cdot \phi_c}{1 - u_6^{S(n)} \cdot c_c} - (\theta - 1) \cdot A_6 \right] \]
\[ B_7^{S(n)} = + \left[ \frac{u_7^{S(h)} \cdot \phi_d}{1 - u_7^{S(h)} \cdot c_d} - (\theta - 1) \cdot A_7 \right] \]
\[ B_8^{S(n)} = + \left[ \frac{1}{2} + \frac{u_8^{S(n)} \cdot \phi_\lambda}{1 - u_8^{S(n)} \cdot c_\lambda} - (\theta - 1) \cdot A_8 \right] \]

C MCMC Estimation

Parameters of interest are
\[ \Theta = (\Theta_Y, \Theta_V) \]
where
\[ \Theta_Y \equiv \begin{pmatrix} \gamma & \psi & \delta & \rho_\lambda & \sigma_\lambda & \mu_c & \mu_d & \mu_\pi \\ \Phi_{c\pi} & \Phi_{d\pi} & \Phi_d & \rho_x & \rho_\pi & \Phi_{x\pi} & \pi_{dc} \end{pmatrix} \]
\[ \Theta_V \equiv \begin{pmatrix} \phi_i & \mu_{h,i} & \sigma_{h,i} \end{pmatrix}_{i \in \{x, \pi, c, d, \iota\}} \]
Observable variables in the model are,

\[ y_{t+1} = \begin{pmatrix} \Delta c_{t+1}, \Delta d_{t+1}, \pi_{t+1}, z_{d,t}, \{ y_h^{(b)} \}_{h \in H} \\ \pi_{t+1} \end{pmatrix} \]

The more detailed state space system is written as

\begin{align*}
    y_{t+1}^o &= A_{t+1} \cdot y_{t+1} \\
    y_{t+1} &= C + H_X \cdot X_{t+1} + H_V \cdot V_{t+1} + \Gamma \cdot e_{t+1} \\
    X_{t+1} &= F_X \cdot X_t + R(V_t) \cdot \varepsilon_{t+1} \\
    V_{t+1} &\sim F_V(V_t; \Theta_V) \quad (32)
\end{align*}

where \( A_{t+1} \) is a selection matrix that takes account of deterministic changes in data availabilities. The vector of stochastic volatilities \( V_{t+1} \) is defined as

\[ V_{t+1} = \begin{bmatrix} \sigma_{x,t+1}^2 & \sigma_{\pi,t+1}^2 & \sigma_{c,t+1}^2 & \sigma_{d,t+1}^2 & \sigma_{t,t+1}^2 \end{bmatrix} \]

such that

\[ \sigma_{i,t+1}^2 \sim \text{Gamma}(\nu_i + \varsigma_{i,t+1}, c_i) \]
\[ \varsigma_{i,t+1} \sim \text{Poisson}\left(\phi_i \cdot \frac{\sigma_{i,t}^2}{c_i}\right) \]

Other variables from the state space system above are,

\[ X_{t+1} \equiv \begin{bmatrix} \chi_{t+1} \\
\eta_{t+1} \\
\chi_t \\
\phi_t \\
\phi_{t-1} \\
\phi_{t-2} \\
\phi_{t-3} \end{bmatrix}, \quad \varepsilon_{t+1} = \begin{bmatrix} \varepsilon^\Lambda_{t+1} \\
\varepsilon^\pi_{t+1} \\
\varepsilon^c_{t+1} \\
\varepsilon^d_{t+1} \\
\varepsilon^\pi_{t+1} \\
\varepsilon^\Lambda_{t+1} \\
\varepsilon^c_{t+1} \end{bmatrix}, \quad \eta_{t+1} \equiv \begin{bmatrix} \sigma^\Lambda \cdot \varepsilon^\Lambda_{t+1} \\
\sigma^\pi \cdot \varepsilon^\pi_{t+1} \\
\sigma^c \cdot \varepsilon^c_{t+1} \\
\sigma^d \cdot \varepsilon^d_{t+1} \\
\sigma^\pi \cdot \varepsilon^\pi_{t+1} \\
\sigma^\Lambda \cdot \varepsilon^\Lambda_{t+1} \end{bmatrix}, \quad \phi_t = \begin{bmatrix} \chi_{c,t-1} \\
\sigma^\Lambda \cdot \varepsilon^\Lambda_t \\
\sigma^{\pi}_{t-1} \cdot \varepsilon^\pi_t \\
\sigma_{c,t-1} \cdot \varepsilon^c_t \end{bmatrix} \]

Given the non-linear nature of the state space model \((32)\), it is difficult to directly evaluate \( P(y \mid \Theta) \). Instead, I rely on the Gibbs sampling method to characterize the joint distribution of

\[ P(\Theta, X, V \mid Y) \]
by iteratively sampling $\Theta$, $X$, and $V$ from relevant conditional distributions. Applying this procedure $J$ yields a large sample of draws,

$$ \left( \Theta^{(j)}, X^{(j)}, V^{(j)} \right)_{j=1}^{J} $$

whose empirical distribution converges to the true theoretical distribution.

Without the loss of generality, assume that after iteration $(g - 1)$, we have sample of $\Theta^{(g-1)}$, $X^{(g-1)}$ and $V^{(g-1)}$.

To move from iteration $(g - 1)$ to iteration $(g)$, I proceed in the following three steps\(^8\)

1. Sample $\Theta^{(g)}$ from $P(\Theta | V^{(g-1)}, Y)$
2. Sample $X^{(g)}$ from $P(X | \Theta^{(g)}, V^{(g-1)}, Y)$
3. Sample $V^{(g)}$ from $P(V | \Theta^{(g)}, X^{(g)}, Y)$

The remaining subsections describe in detail the implementation of each step.

### C.1 Sample $\Theta$ from $P(\Theta | V, Y)$

I use random-walk Metropolis Hastings to accept/reject new draws of $\Theta$. The relevant posterior is

$$ P(\Theta | V, Y) \propto P(\Theta) \cdot P(V | \Theta) \cdot P(Y | \Theta, V) $$

Given $\Theta^{(g-1)}$ from the previous iteration, I draw $\tilde{\Theta}$ from

$$ \mathcal{N}(\Theta^{(g-1)}, \Sigma) $$

and set

$$ \Theta^{(g)} = \begin{cases} 
\Theta^{(g-1)} & \text{w.p. } \alpha \\
\tilde{\Theta} & \text{w.p. } (1 - \alpha)
\end{cases} $$

where

$$ \alpha = \min \left\{ 1, \frac{P(\tilde{\Theta} | V, Y)}{P(\Theta^{(g-1)} | V, Y)} \right\}. $$

The variance covariance matrix of the proposal distribution is fine tuned so that about 25% of proposed draws are accepted.

---

\(^8\)Strictly speaking, these three steps form a proper Gibbs sampler with two blocks, $(\Theta, X)$ and $V$, where steps 1 and 2 can be considered two sub-steps within a larger block step that samples $(\Theta, X)$ from $P(X, \Theta | V, Y)$. 

43
C.2 Sample $X$ from $P(X \mid \Theta, V, Y)$

Conditional on the history of $V$, we have a linear Gaussian state space model,

$$y_{t+1}^o = A_{t+1} \cdot \left( C + H_X \cdot X_{t+1} + H_V \cdot V_{t+1} + \Sigma_e \cdot e_{t+1} \right)$$

$$X_{t+1} = F_X \cdot X_t + R(V_t) \cdot \varepsilon_{t+1}.$$

Thus we can draw $X$ using the standard forward filtering, and backward sampling method discussed by Carter and Kohn (1994) and Kim, Nelson, et al. (1999). From this system, we can get for all $t = 1, \cdots, T$ the filtered estimates of $X$ $X_{t|t} \equiv E(X_t \mid I_t)$, $P_{t|t} \equiv V(X_t \mid I_t)$.

Note that the VCV of innovation matrix

$$Q_t \equiv R(V_t) \cdot R(V_t)'$$

is singular. To draw $\chi^T$ and $\eta^T$ from normal distributions with mean and variance,

$$E(X_t \mid I_T), \quad V(X_t \mid I_T).$$

I run the backward smoother on the transition equation of $X_{t+1}^* = [\chi_{t+1}, \eta_{t+1}]$, such that

$$\begin{bmatrix} \chi_{t+1} \\ \eta_{t+1} \end{bmatrix}_{X_{t+1}} = \begin{bmatrix} F_X \\ \sigma_{X} \end{bmatrix}_{X^*} \begin{bmatrix} \chi_t \\ \eta_t \end{bmatrix}_{X_t} + \begin{bmatrix} R_X(V_t) \\ \sigma_{R} \end{bmatrix}_{V_t} \begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \\ \epsilon_{t+1} \end{bmatrix}_{\epsilon_{t+1}}.$$

Note that in this transition equation, the VCV for innovation terms are non-singular,

$$Q_t^* \equiv R^* (V_t) \cdot R^* (V_t)'$$
C.3 Sampling V Conditional on \((\Theta, X)\)

C.3.1 Setup of problem

To reduce the number of particles needed, I sequentially draw each one of the element series of \(V\), denoted as \(V_i\), conditional on the values of other elements of \(V\), denoted as \(V_{\bar{i}}\).

From the observation equation of asset prices \(\text{[13]}\), and conditional on \(V_{\bar{i}}\), I define

\[
\hat{p}_{i,t+1} \equiv p_{t+1} - \left( \Psi_0 + \Psi_{\chi} \cdot \chi_{t+1} + \sum_{j \neq i} \Psi_{V,j} \cdot V_{j,t+1} \right),
\]

for \(i \in \{x, \pi, c, d, \iota\}\).

Thus, when conditional on values of \(X^*_{t+1} \equiv [\chi_{t+1}, \eta_{t+1}]\) and \(\Theta\), we only need to draw \(V_i \equiv \{\sigma_{i,t}^2\}\) from the non-linear state space system below;

\[
\begin{align*}
\hat{P}_{i,t+1} &= \Psi_{V,i} \cdot \sigma_{i,t+1}^2 + \Sigma_p \cdot e_{t+1}^P \\
\hat{z}_{i,t+1} &= \sigma_{i,t} \cdot \varepsilon_{i,t+1} \\
\sigma_{i,t+1}^2 &\sim \text{Gamma} \left( \nu_i + \varsigma_{i,t+1}, c_i \right) \\
\varsigma_{i,t+1} &\sim \text{Poisson} \left( \phi_i \cdot \frac{\sigma_{i,t+1}^2}{c_i} \right)
\end{align*}
\]

C.3.2 Particle Gibbs Sampler

For the rest of this section, we omit \(\Theta\) and \(\chi\), and focus on inference of the unobserved sequence of \(S\) condition on the sequence of observables \(Y\). Suppose from the previous iteration, that we have draws of hidden state \(S^{(g-1)}\). The PG sampler consists of two main steps. Let \(J\) be total number of particles.

- **Step 1.** Forward simulation. We generate for each \(t = 0, 1, \cdots, T\), a set of \(J\) candidates values \(\left\{ S_t^{(j)} \right\}_{j=1}^J\), and its importance weight \(\left\{ \omega_t^{(j)} \right\}_{j=1}^J\).

- **Step 2.** Backward selection. We sample from time \(T\) back to time 0 a set of \(S_t^{(g)}\) based on importance weight.

**Forward simulation.** For \(t = 1, \cdots, T\), run:

- For \(j = 1\), set \(S_t^{(1)} = S_t^{(g-1)}\).
- For \(j = 2, \ldots, J\), draw from proposal \(q \left( S_t^{(j)} \mid S_t^{(j)} \right)\).
- For \(j = 1, \ldots, J\), calculate importance weight,

\[
\omega_t^{(j)} = \frac{P \left( Y_t \mid S_t^{(j)}, S_{t-1}^{(j)} \right) \cdot P \left( S_t^{(j)} \mid S_{t-1}^{(j)} \right)}{q \left( S_t^{(j)} \mid S_{t-1}^{(j)} \right)}
\]
• For \( j = 1, \ldots, J \), normalize weights

\[
\hat{\omega}_t^{(j)} = \frac{\omega_t^{(j)}}{\sum_{j=1}^{J} \omega_t^{(j)}}
\]

• Re-sample to generate a new sample of \( \{S_t^{(j)}\}_{j=1}^{J} \). We do so by setting \( S_t^{(1)} = S_t^{(g-1)} \), and sample \((J - 1)\) times from distribution

\[
\{S_t^{(j)}, \hat{\omega}_t^{(j)}\}_{j=1}^{J}.
\]

We specify that the proposal distribution of \( q(S_t \mid S_{t-1}) \) is the same as the transition density \( P(S_t \mid S_{t-1}) \). Thus, to it is straightforward to draw from \( q(S_t \mid S_{t-1}) \), by first drawing \( \varsigma_t \) from the Poisson distribution with \( \lambda = \phi(S_{t-1}/c) \), and then drawing \( S_t \) from Gamma \((\nu + \varsigma_t, c)\).

This choice would also imply that we can simplify the expression for importance weight,

\[
\omega_t^{(j)} = P(Y_t \mid S_t^{(j)}, S_{t-1}^{(j)}) \cdot P(S_t \mid S_{t-1}) = P(Y_t \mid S_t^{(j)}, S_{t-1}^{(j)})
\]

\[
\omega_t^{(j)} \propto f_N(\hat{P}_{i,t+1}; \Psi_S S_t^{(j)}, \Sigma_p \Sigma_p') \cdot f_N(\hat{z}_{i,t+1}; 0, [S_{t-1}^{(j)}]^2)
\]

where \( f_N(x \mid \mu, \Omega) \) is the density of normal distribution with mean \( \mu \) and variance covariance matrix \( \Omega \).

**Backwards sampling.** From forward simulation, we have stored \( \{S_t^{(j)}, \hat{\omega}_t^{(j)}\}_{t=1, j=1}^{T,J} \)

• At time \( t = T \), we draw \( S_T^* = S_T^{(j)} \) with probability \( \hat{\omega}_T^{(j)} \).

• At time \( t = T - 1, \ldots, 0 \), run:

  - For \( j = 1, \ldots, J \), calculate backwards weights

\[
\omega_{t|T}^{(j)} \propto \hat{\omega}_t^{(j)} \cdot P(S_{t+1}^* \mid S_t^{(j)})
\]

  - For \( j = 1, \ldots, J \), normalize weights

\[
\hat{\omega}_{t|T}^{(j)} = \frac{\omega_{t|T}^{(j)}}{\sum_{j=1}^{J} \omega_{t|T}^{(j)}}
\]

  - Draw a particle \( S_t^* = S_t^{(j)} \) with probability \( \hat{\omega}_t^{(j)} \)

• Set \( S^{(g)} = S^* \).
The goal of this part is to show that, conditional on values of $V$, our model can be cast as

$$
\begin{align*}
    y_{t+1} & = A_{t+1} \cdot y_{t+1} = A_{t+1} \cdot (C + H_X \cdot X_{t+1} + H_V \cdot V_{t+1} + \Sigma_e \cdot e_{t+1}) \\
    X_{t+1} & = F_X \cdot X_t + R(V_t) \cdot \varepsilon_{t+1}
\end{align*}
$$

The model is specified at monthly frequency. All variables are observed at quarterly frequency. Inflation and dividend growth are modeled to be observed without measurement error. Monthly consumption series $\Delta c_{o t+1}$ are observed with measurement error, while quarterly consumption growth rates $\Delta c_{q t+1}$ are assumed to be observed without measurement error, at the end-of-quarterly months. It follows that

$$
\Delta c_{q t+1} = \sum_{j=1}^{5} \omega_j \Delta c_{t+1-j+1}
$$

where

$$
\omega_j = \left( \frac{3 - |j - 3|}{3} \right), \quad \text{for } j = 1, \cdots, 5.
$$

Thus, the quarterly consumption growth rate is a ‘tent-shaped’ weighted average of past consumption growth. Let $A_t$ be deterministic selection matrix that determines which variables are observable at time $t$. At end-of-quarter month, $z_t = [\Delta c^q_t, \Delta c^q_d, \pi_t]$; otherwise $z_t = [\Delta c^o_t \Delta d_t, \pi_t]$. We stack economic fundamentals and asset prices together to form our observable variable that $y_{t+1} = [z_{t+1}, p_{t+1}]$. Finally we have the following state space representation which takes into consideration data availability,

$$
\begin{align*}
    y_{t+1} & = A_{t+1} \cdot y_{t+1} = A_{t+1} \cdot (C + H_X \cdot X_{t+1} + H_V \cdot V_{t+1} + \Sigma_e \cdot e_{t+1}) \\
    X_{t+1} & = F_X \cdot X_t + R(V_t) \cdot \varepsilon_{t+1}
\end{align*}
$$

Asset prices $p_{t+1}$ are affine functions of state variables such that

$$
p_{t+1} = \Psi_0 + \Psi_S \cdot \begin{bmatrix} \chi_{t+1} \\
\eta_{t+1} \\
\chi_{t+1} \end{bmatrix} = \Psi_0 + \Psi_X \cdot \chi_{t+1} + \Psi_V \cdot V_{t+1}
$$

where state variables $X_{t+1}$ are defined as,

$$
X_{t+1} = \begin{bmatrix} \chi_{t+1} \\ \eta_{t+1} \\ \chi_{t} \end{bmatrix}
$$
The transition equation for \( \chi_{t+1} \) is defined as the follows:

\[
\begin{bmatrix}
  x_{\Lambda,t+1} \\
  x_{c,t+1} \\
  x_{\pi,t+1}
\end{bmatrix}
= 
\begin{bmatrix}
  \rho_\Lambda & 0 & 0 \\
  0 & \rho_x & 0 \\
  0 & 0 & \rho_\pi
\end{bmatrix}
\begin{bmatrix}
  x_{\Lambda,t} \\
  x_{c,t} \\
  x_{\pi,t}
\end{bmatrix}
+ 
\begin{bmatrix}
  \sigma_\Lambda & 0 & 0 \\
  0 & \sigma_{x,t} & \Phi_{x\pi} \cdot \sigma_{\pi t} \\
  0 & 0 & \sigma_{\pi t}
\end{bmatrix}
\begin{bmatrix}
  \varepsilon_{\Lambda,t+1} \\
  \varepsilon_{x,t+1} \\
  \varepsilon_{\pi t+1}
\end{bmatrix}
\]

Given this, the state transition equation for extended system

\[
\begin{bmatrix}
  \chi_{t+1} \\
  \eta_{t+1} \\
  \chi_t
\end{bmatrix}
= 
\begin{bmatrix}
  \mathbf{F}_\chi & 0 & 0 \\
  0 & \mathbf{I}_3 & 0 \\
  \mathbf{I}_3 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
  \chi_t \\
  \eta_t \\
  \chi_{t-1}
\end{bmatrix}
+ 
\begin{bmatrix}
  \mathbf{R}_\chi (V_t) & 0 & 0 \\
  0 & \mathbf{R}_\eta (V_t) & 0 \\
  0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
  \varepsilon_{\chi,t+1} \\
  \varepsilon_{\eta,t+1} \\
  \varepsilon_{t+1}
\end{bmatrix}
\]

Denote the subset of state variables that are relevant for determination of \( \Delta c_{t+1} \) as

\[
\phi_{t+1} = 
\begin{bmatrix}
  x_{c,t} \\
  \sigma_\Lambda \cdot \varepsilon_{\Lambda,t+1} \\
  \sigma_{\pi,t} \cdot \varepsilon_{\pi,t+1} \\
  \sigma_{c,t} \cdot \varepsilon_{c,t+1} \\
  \sigma_{d,t} \cdot \varepsilon_{d,t+1} \\
  \sigma_{\pi,t} \cdot \varepsilon_{\pi,t+1} \\
  \sigma_{c,t} \cdot \varepsilon_{c,t+1}
\end{bmatrix}
\]

Recall the observation equation for monthly consumption such that

\[
\Delta c_{t+1} = \mu_c + x_{c,t} + \pi_{c\Lambda} \cdot \varepsilon_{\Lambda,t+1} + \Phi_{x\pi} \cdot \sigma_{x,t} \cdot \varepsilon_{\pi,t+1} + \sigma_{c,t} \cdot \varepsilon_{c,t+1}
\]

The state variables for this process are \( \phi_{t+1} \) where

\[
\Delta c_{t+1} \equiv \mu_c + h_c \cdot \phi_{t+1} = \mu_c + \left[ \begin{array}{ccc}
  \frac{\pi_{c\Lambda}}{\sigma_\Lambda} & \Phi_{x\pi} & 1
\end{array} \right] \cdot \phi_{t+1}
\]
Finally, we can rewrite the system as follows:

\[
\begin{bmatrix}
\chi_{t+1} \\
\eta_{t+1} \\
\chi_t \\
\phi_t \\
\phi_{t-1} \\
\phi_{t-2} \\
\phi_{t-3}
\end{bmatrix}
= F_\phi \cdot
\begin{bmatrix}
\chi_t \\
\eta_t \\
\chi_{t-1} \\
\phi_{t-1} \\
\phi_{t-2} \\
\phi_{t-3} \\
\phi_{t-4}
\end{bmatrix}
+ R_X(V_t) \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]

We then can write the following:

\[
\begin{bmatrix}
\Delta c_{t+1}^e \\
\Delta c_{t+1} \\
\Delta d_{t+1} \\
\pi_{t+1} \\
\{p_{i,t+1}\}_i \\
\{\bar{p}_{i}\}_i
\end{bmatrix}
= 3 \cdot \mu_c 
\begin{bmatrix}
0 \\
\mu_c \\
\mu_d \\
\mu_\pi \\
\{\pi_i\}_i \\
\{\bar{\pi}_i\}_i
\end{bmatrix}
\begin{bmatrix}
\sigma_{x,t+1} \\
\sigma_{\pi,t+1} \\
\sigma_{d,t+1} \\
\sigma_{x,t+1} \\
\sigma_{\pi,t+1} \\
\sigma_{d,t+1}
\end{bmatrix}
+ \Sigma_e \cdot e_{t+1}
\]

\[
= \begin{bmatrix}
[\omega_1 \cdot \left( \frac{\pi c}{\sigma_A} \right)] \\
[\omega_1 \Phi c x] \\
[\omega_1 \Phi d x] \\
[\omega_1 \Phi c x] \\
[\omega_1 \Phi d x] \\
[\omega_1 \Phi c x]
\end{bmatrix} + H_X
\]

\[
\begin{bmatrix}
\chi_{t+1} \\
\sigma_{x,t+1} \\
\sigma_{\pi,t+1} \\
\sigma_{d,t+1} \\
x_{c,t} \\
x_{\pi,t} \\
\phi_t \\
\phi_{t-1} \\
\phi_{t-2} \\
\phi_{t-3}
\end{bmatrix}
\]
where $\omega \equiv [\omega_2 \omega_3 \omega_4 \omega_5]$.

E Solution of the Affine Term Structure Model

Model setup:

$$
X_t = \mu + \Phi \cdot X_{t-1} + \Sigma \cdot \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I)
$$

$$
r_t = \delta_0 + \delta_1' \cdot X_t
$$

$$
\lambda_t = \lambda_0 + \lambda_1 \cdot X_t
$$

$$
m_{t+1} \equiv \log (M_{t+1}) = -\frac{1}{2} \lambda_t' \lambda_t - r_t - \lambda_t' \cdot \varepsilon_{t+1}
$$

Non-arbitrage implies:

$$
P_t^{(n+1)} = E_t \left( M_{t+1} \cdot P_t^{(n)} \right)
$$

$$
p_t^{(n)} \equiv \log \left[ P_t^{(n+1)} \right] = \overline{A}_n + \overline{B}_n' \cdot X_t.
$$

Then bond yields follows as

$$
y_t^{(n)} = A_n + B_n' \cdot X_t
$$

where $A_n = -\frac{1}{n} \cdot \overline{A}_n$ and $B_n = -\frac{1}{n} \cdot \overline{B}_n$.

- Base case ($n = 1$). Recall that $P_t^{(1)} = E_t (M_{t+1} \cdot 1)$, it is obvious that

$$
p_t^{(1)} = -r_t = -\delta_0 - \delta_1' \cdot X_t
$$

so that $\overline{A}_1 = -\delta_0$ and $\overline{B}_1 = -\delta_1$.

- Induction case ($n \geq 1$). Given that $p_t^{(n)} = \overline{A}_n + \overline{B}_n \cdot X_t$, by no-arbitrage condition,

$$
P_t^{(n+1)} = E_t \left( M_{t+1} \cdot P_t^{(n)} \right)
$$

It can be shown that

$$
p_t^{(n+1)} \equiv \log P_t^{(n+1)} = \log E_t \left[ \exp \left( m_{t+1} + p_t^{(n)} \right) \right]
$$

$$
= \log E_t \left\{ \exp \left[ \left( -\frac{1}{2} \lambda_t' \lambda_t - r_t - \lambda_t' \cdot \varepsilon_{t+1} \right) + \left( \overline{A}_n + \overline{B}_n' \cdot X_{t+1} \right) \right] \right\}
$$

$$
= \left( -\delta_0 + \overline{A}_n + \overline{B}_n' \mu + \frac{1}{2} \overline{B}_n' \Sigma \Sigma' \overline{B}_n - \overline{B}_n' \Sigma \lambda_0 \right) + \left( \overline{B}_n' \Phi - \delta_1' - \overline{B}_n' \Sigma \lambda_1 \right) \cdot X_t
$$

Thus,

$$
\overline{A}_{n+1} = -\delta_0 + \overline{A}_n + \overline{B}_n' \left( \mu - \Sigma \lambda_0 \right) + \frac{1}{2} \overline{B}_n' \Sigma \Sigma' \overline{B}_n
$$

$$
\overline{B}_{n+1} = -\delta_1' + \overline{B}_n' \left( \Phi - \Sigma \lambda_1 \right).
$$