The London School of Economics and Political Science

Essays on the Term Structures of Bonds and Equities

Wen Yan

Thesis submitted to the Department of Finance
of the London School of Economics for the degree of PhD in Finance

July 2015

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 23528 words.

Abstract

Chapter 1 “The Term Structure of Equities” examines the term structure of equities. Using observed prices of dividend strips, prices of zero-coupon equities are extracted, and their yields and returns characteristics are documented. An affine term structure model is used to model the term structure of equities. The model is estimated, and model-implied equity yields and returns are shown to match the data well. However, the model-implied long-run risk-neutral mean of the short rate is implausible. (The next chapter takes this into account and estimates bond and equity yield curves jointly using data on both zero-coupon bonds and zero-coupon equities.)

Chapter 2 “Estimating a Unified Framework of Co-Pricing Stocks and Bonds” estimates a maximal identifiable affine term structure model that explains the joint prices of stocks and bonds. Using the test assets of Treasury bonds and dividend strips, it is shown that the estimated model can generally match the time series and cross-sectional properties of zero-coupon bonds, zero-coupon equities and the aggregate stock index. Moreover, imposing restrictions prevalent in the co-pricing literature on the maximal model enhances certain features of the model such as the high return of the short-term dividend strip, but reduces the model’s ability to fit other aspects of the data such as the level of the

market risk premium.

Chapter 3 "The Role of Asian Countries" Reserve Holdings on the International Yield Curves" studies the effect of Asian countries' reserve holdings on the yield curves of six industrialized countries: the United States, the United Kingdom, Germany, Canada, Switzerland and Australia. A Gaussian affine term structure model with three yield factors and three unspanned macro factors including reserves is estimated to fit the yield curve of each country. Impulse responses and variance decompositions show that Asian countries' reserve holdings are an important factor affecting the international yield curves.

Acknowledgment

I thank Mikhail Chernov, Ian Martin and Philippe Mueller for their invaluable advice and continuous encouragement. I also thank Christian Julliard, Dong Lou, Igor Makarov, Christoper Polk, Andrea Tamoni, Dimitri Vayanos, Andrea Vedolin and other participants at the LSE finance seminars for their constructive comments and helpful suggestions.

Contents

List of Figures

List of Tables

Chapter 1

The Term Structure of Equities

1.1 Abstract

This chapter examines the term structure of equities. Using observed prices of dividend strips, prices of zero-coupon equities are extracted, and their yields and returns characteristics are documented. An affine term structure model is used to model the term structure of equities. The model is estimated, and model-implied equity yields and returns are shown to match the data well. However, the model-implied long-run risk-neutral mean of the short rate is implausible. (The next chapter takes this into account and estimates bond and equity yield curves jointly using data on both zero-coupon bonds and zero-coupon equities.)

1.2 Introduction

There is an extensive literature on identifying the common factors that affect the bond yield curve. The work by Litterman and Scheinkman (1991) showed that three factors, namely "level", "slope" and "curvature", explain over 90% of cross-sectional bond yield variations for almost any reasonable length of sample period and any combination of yield maturities. This result is so robust that factor analysis has since populated the analysis of bond term structure, with a fourth factor coined by Cochrane and Piazzesi (2005), namely "the return forecasting factor", and the discovery of the "hidden factor" by Duffee (2011).

However, despite the advancement of factor analysis on the bond yield curve, the equity yield curve is rarely studied. The literature has mostly focused on studying the risk and return behavior of the aggregate stock market without looking at the individual terms that comprise it. However, because the value of the aggregate stock market can be viewed as the total value of the discounted future dividend payments (Gordon 1962), in addition to studying the aggregate price of these dividend payments, exploring the properties of each dividend payment should provide us with valuable information about the way stock prices are formed and improve our understanding of investors' risk preferences and the endowment or technology process in macro-finance models. Hence in this paper, I study the term structure of equities by exploring the properties of the individual dividend payments that comprise the aggregate stock market. More specifically, I focus on zero-coupon equities, a concept created using the analogy of zero-coupon bonds. Just like a zero-coupon bond giving the investor a fixed payment at the end of the bond's maturity, a zero-coupon equity simply gives the investor a variable payment, that is, the stochastic dividend, at the end of the security's maturity. This stochastic dividend could be paid out by a particular company, industry or the aggregate economy. And the sum of discounted future dividend payments will be the value of the company, industry or the aggregate economy. In this paper, I focus on the term structure of the stochastic dividends of varying maturities paid out by the aggregate stock market index. And the first important questions for us are what the yield and return characteristics of the zero-coupon equities are and whether there exist common factors that can price the equity yield curve well.

Previously, the lack of study on the equity yield curve was largely due to data unavailability. To study the bond yield curve, we use monthly zero-coupon bond yields data dating back to the post war period (or earlier) – these are often available for maturities from one month to 30 years. But such data has not been available for equities. However, my study of the equity term structure is made possible by the availability of dividend strips data on the S&P 500 from van Binsbergen, Brandt and Koijen (2012). Specifically, whereas a zero-coupon equity of maturity n at time t gives the investor a stochastic dividend payment from $t + n - 1$ to $t + n$, buying a dividend strip of maturity n at time t entitles the investor to all the dividends paid out from time t to $t + n$. Using put–call parity and observed put and call prices on the S&P 500, i.e. Long-Term Equity Anticipation Securities (LEAPS) from the Chicago Board Options Exchange (CBOE), van Binsbergen, Brandt and Koijen decompose the index into a long-term equity and a short-term equity, which is the dividend strip. In particular, prices of dividend strips for maturities of six, twelve, eighteen and twenty-four months are priced, and I extract zero-coupon equity prices from the prices of these strips.

The model used to consistently price dividends will follow Lettau and Wachter (2007), which is an extension of the bond affine term structure models first proposed by Duffie and Kan (1996). The economy is driven by three state variables. One is the short rate factor, one is the growth rate of dividend and the other is the first principal component of the set of zero-coupon equities, which can be interpreted as a portfolio of equity yields. The inclusion of the growth rate of dividend and the short rate is due to the fact that these two variables are crucial components in the pricing model of zero-coupon equities. Regarding the estimation of the model, identification is ensured by showing that the three-factor model is observationally equivalent to the maximally identified canonical model specified in terms of three entirely latent factors – a method developed in Joslin, Singleton and Zhu (2011).

By estimating the term structure of equities, this paper fills the gap in the literature in which only the term structure of bonds was estimated. Both Lettau and Wachter (2007) and Lettau and Wachter (2011) study the equity term structure, but they do not use data on zero-coupon equities or dividend strips. Moreover, what we learn using an estimated factor model is that we can see from the data the dynamics among factors. The calibration exercise usually makes strong and sometimes counterfactual assumptions on the dynamics of factors and the interaction between them. For example, in Lettau and Wachter (2011), dividend growth follows an autoregressive process with positive autocorrelation coefficient. However, in the data, dividend growth is strongly negatively autocorrelated. Such restrictions will likely distort the model's predictions of asset prices and each factor's implication on the asset prices.

The rest of the chapter is structured as follows: Section 1.3 outlines the affine term structure model that is able to price zero-coupon equities, the aggregate stock market index as well as zero-coupon bonds. Section 1.4 describes the data, estimation strategy and estimation results and shows the model implications for bond yields. Section 1.5 concludes.

1.3 The model

This section introduces the general Gaussian affine term structure model that is able to price both bonds and equities. The model follows the affine framework first proposed by Duffie and Kan (1996) to price zero-coupon bonds with different maturities, which has been further extended in the literature to price zero-coupon equities with different maturities. Affine term structure models have been widely used in the bond pricing literature mainly due to their ability to generate tractable solutions for bond yields. The same benefit of tractability can be carried forward to equity pricing. And the way of using affine term structure techniques to value zero-coupon equities for each maturity and summing over the prices of zero-coupon equities of all maturities to reach the aggregate market index has been applied in Ang and Liu (2004), Bekaert, Engstrom and Grenadier (2010), Lettau and Wachter (2007), Lettau and Wachter (2011) and Wachter (2006).

1.3.1 The economy

It is assumed that the economy at time t is driven by a state vector X_t that follows a VAR(1) process under both the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$,

$$\Delta X_t = K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}}, \tag{1.1}$$

$$\Delta X_t = K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, \tag{1.2}$$

where X_t is an $N \times 1$ vector, $K_{0X}^{\mathbb{P}}$ and $K_{0X}^{\mathbb{Q}}$ are $N \times 1$ vectors, $K_{1X}^{\mathbb{P}}$, $K_{1X}^{\mathbb{Q}}$ and Σ_X are $N \times N$ matrices and both $\epsilon_t^{\mathbb{P}}$ and $\epsilon_t^{\mathbb{Q}}$ are $N \times 1$ vectors of independent shocks to various risk factors affecting the economy with mean zero and unit variance.

Let $r_t = \log R_t$, the one-period interest rate, be an affine function of the state vector,

$$r_t = \rho_{0X} + \rho_{1X}' X_t, \tag{1.3}$$

where ρ_{0X} is a scalar and ρ_{1X} is an $N \times 1$ vector.

The level of the aggregate dividend of the economy is denoted by D_t. Let $d_t = \log D_t$, and the log dividend growth rate from time $t-1$ to time t be defined as $\Delta d_t = \log(D_t/D_{t-1})$. To maintain the affine structure of the model, the dividend growth process is assumed to be an affine function of the underlying state vector, i.e.

$$\Delta d_t = \delta_{0X} + \delta'_{1X} X_t, \tag{1.4}$$

where δ_{0X} is a scalar and δ_{1X} is an $N \times 1$ vector.

Let the market price of risk vector λ_t be affine in the state vector,

$$\lambda_t = \lambda_{0X} + \lambda_{1X} X_t. \tag{1.5}$$

Here λ_t is an $N \times 1$ vector of time-varying market prices of risk, λ_{0X} is an $N \times 1$ vector and λ_{1X} is an $N \times N$ matrix.

By no-arbitrage, we obtain the pricing kernel or the stochastic discount factor (SDF) M_{t+1} of the economy as

$$M_{t+1} = \exp(-r_t - \tfrac{1}{2}\lambda'_t \lambda_t - \lambda'_t \epsilon_{t+1}), \tag{1.6}$$

which can be used to consistently price all assets. That is, we have the Euler equation

$$1 = E_t[M_{t+1} R_{t+1}], \tag{1.7}$$

where R_{t+1} is the one-period return on any asset in the economy.

1.3.2 Zero-coupon equities

To price equities, I follow the approach in, for example, Lettau and Wachter (2007) for defining the zero-coupon equity, which is analogous to the concept of zero-coupon bond. It is assumed that the zero-coupon equity is an asset that pays off the aggregate dividend at some fixed maturity, and its price is an exponential affine function of the underlying state vector. As a result, the price of zero-coupon equity has an analytical form that is similar to the price of the zero-coupon bond.

Specifically, let P^{d}_{nt} denote the time-t price of a zero-coupon equity of maturity n, that is, the time-t price of the aggregate dividend paid out between time $t+n-1$ and time $t+n$. This implies that its one-period return from t to $t+1$ can be written as

$$R^{\mathrm{d}}_{n,t+1} = \frac{P^{\mathrm{d}}_{n-1,t+1}}{P^{\mathrm{d}}_{n,t}} = \frac{P^{\mathrm{d}}_{n-1,t+1}/D_{t+1}}{P^{\mathrm{d}}_{n,t}/D_t}\frac{D_{t+1}}{D_t}. \tag{1.8}$$

Plugging this into the Euler equation implies that the price scaled by the aggregate dividend will satisfy the following equation:

$$\frac{P^{\mathrm{d}}_{nt}}{D_t} = E_t\left[M_{t+1}\frac{P^{\mathrm{d}}_{n-1,t+1}}{D_{t+1}}\frac{D_{t+1}}{D_t}\right]. \tag{1.9}$$

If we write the scaled equity price as an exponential affine function of the state vector, i.e.

$$\frac{P^{\mathrm{d}}_{nt}}{D_t} = \exp(A^{\mathrm{d}}_n + B^{\mathrm{d}\prime}_n X_t), \tag{1.10}$$

then all quantities in the Euler equation can now be expressed as exponential affine functions of the state vector. Moreover, using the Euler equation, we can express the constant A^{d}_n and the $1 \times N$ loadings of the scaled equity price on the state vector, i.e. $B^{\mathrm{d}\prime}_n$, as functions of the underlying parameters of the model by solving a set of Riccati equations with the boundary condition $P^{\mathrm{d}}_{0t}/D_t = 1$.

More specifically, the loadings are solved recursively as follows:

$$\begin{aligned} A^{\mathrm{d}}_n &= -(\rho_{0X} - \delta_{0X}) + A^{\mathrm{d}}_{n-1} + (\delta_{1X} + B^{\mathrm{d}}_{n-1})' K^{\mathbb{Q}}_{0X} \\ &\quad + \tfrac{1}{2}(\delta_{1X} + B^{\mathrm{d}}_{n-1})' \Sigma_X \Sigma_X' (\delta_{1X} + B^{\mathrm{d}}_{n-1}), \end{aligned} \tag{1.11}$$

$$B^{\mathrm{d}\prime}_n = -\rho_{1X}' + (\delta_{1X} + B^{\mathrm{d}}_{n-1})'(K^{\mathbb{Q}}_{1X} + I), \tag{1.12}$$

with the starting values of $A_0^{\mathrm{d}} = 0$ and $B_0^{\mathrm{d}} = 0$. Details of the derivation are provided in Appendix A1.1.

Moreover, we have

$$y_{nt}^{\mathrm{d}} = -\frac{1}{n} \ln \frac{P_{nt}^{\mathrm{d}}}{D_t} = -\frac{1}{n}(A_n^{\mathrm{d}} + B_n^{\mathrm{d}\prime} X_t) = -\frac{1}{n} A_n^{\mathrm{d}} - \frac{1}{n} B_n^{\mathrm{d}\prime} X_t. \tag{1.13}$$

When we use the affine model to price bonds, bond yields are affine functions of the state vector, and in estimation we try to match the estimated bond yields with the observed bond yields. Analogously, when we use the affine model to price equities, the quantity

$$y_{nt}^{\mathrm{d}} = -\frac{1}{n} \ln \frac{P_{nt}^{\mathrm{d}}}{D_t}$$

will be the equity "yield" that we try to match.

1.3.3 The aggregate market

Since a zero-coupon equity is an asset that pays off the aggregate dividend at some fixed maturity, by summing the prices of zero-coupon equities of all maturities, we get the aggregate market index. Note here that, unlike the prices of zero-coupon equities, which are exponential affine functions of the state vector, the market index will not be an exponential affine function of the state vector

$$P_t^{\mathrm{m}} = \sum_{n=1}^{\infty} P_{nt}^{\mathrm{d}} = \sum_{n=1}^{\infty} \exp(A_n^{\mathrm{d}} + B_n^{\mathrm{d}\prime} X_t) \times D_t. \tag{1.14}$$

1.3.4 Zero-coupon bonds

Although zero-coupon bonds are not studied in this chapter, they will be studied jointly with zero-coupon equities in Chapter 2. Therefore, for completeness and convenience, the pricing equations of zero-coupon bonds are outlined here. To price nominal bonds, let P^{b}_{nt} denote the time-t price of the n-period nominal zero-coupon bond. Assuming bond price is exponential affine in the state vector

$$P^{\mathrm{b}}_{nt} = \exp(A^{\mathrm{b}}_n + B^{\mathrm{b}\prime}_n X_t), \tag{1.15}$$

where A^{b}_n is a scalar and B^{b}_n is an $N \times 1$ vector, by the Euler equation we have

$$P^{\mathrm{b}}_{nt} = E_t[M_{t+1} P^{\mathrm{b}}_{n-1,t+1}] \tag{1.16}$$

with the boundary condition $P^{\mathrm{b}}_{0t} = 1$.

The price of the zero-coupon bond $P^{\mathrm{b}}_{nt} = \exp(A^{\mathrm{b}}_n + B^{\mathrm{b}\prime}_n X_t)$ has exactly the same form as the scaled price of the zero-coupon equity, which is $P^{\mathrm{d}}_{nt}/D_t = \exp(A^{\mathrm{d}}_n + B^{\mathrm{d}\prime}_n X_t)$. And the Euler equation for the zero-coupon bond is exactly the same as the Euler equation for the zero-coupon equity without the dividend growth process. Hence we can just take the results from the zero-coupon equity, set $\delta_{0X} = 0$ and $\delta_{1X} = 0$ and change the superscript from d to b to get the standard solutions for bond prices' loadings on the state vector, which are

$$A^{\mathrm{b}}_n = -\rho_{0X} + A^{\mathrm{b}}_{n-1} + {B^{\mathrm{b}}_{n-1}}' K^{\mathbb{Q}}_{0X} + \tfrac{1}{2} {B^{\mathrm{b}}_{n-1}}' \Sigma_X \Sigma'_X B^{\mathrm{b}}_{n-1}, \tag{1.17}$$

$$B^{\mathrm{b}\prime}_n = -\rho'_{1X} + {B^{\mathrm{b}}_{n-1}}'(K^{\mathbb{Q}}_{1X} + I), \tag{1.18}$$

with the starting values being $A^{\mathrm{b}}_n = 0$ and $B^{\mathrm{b}}_n = 0$.

And bond yield of maturity n can be expressed as

$$y^{\mathrm{b}}_{nt} = -\frac{1}{n} \ln P^{\mathrm{b}}_{nt} = -\frac{1}{n}(A^{\mathrm{b}}_n + B^{\mathrm{b}\prime}_n X_t) = -\frac{1}{n} A^{\mathrm{b}}_n - \frac{1}{n} B^{\mathrm{b}\prime}_n X_t. \tag{1.19}$$

1.4 Estimation

1.4.1 The normalized model

The general Gaussian dynamic term structure model previously stated in Section 1.3 is not ready to be estimated, as any affine transformations of the state process are observationally equivalent as shown in Dai and Singleton (2000). That is, suppose we have two models; they are the same in all aspects except that the state process in one model is an affine transformation of the state process in the other model. Then the two models will generate exactly the same asset prices. In other words, given the same set of asset prices, there exists an infinite number of models that can generate this set of asset prices. Therefore, these models are generally not identified without imposing restrictions. To impose the minimum number of restrictions on the state process such that the model is identified, we can follow Joslin, Singleton and Zhu (2011) (JSZ) and normalize all models to a canonical form, in which the state vector X_t is entirely latent. As a result of this normalization, given the same set of asset prices, there is only one model in the canonical form that can generate this set of asset prices. The JSZ canonical form is only able to price zero-coupon bonds. I extend the canonical form in JSZ to a canonical form that is able to price zero-coupon equities as well.

Let us first recall the general form of the state process in Section 1.3, in which the parameters are unrestricted:

$$\begin{aligned}
\Delta X_t &= K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}}, \\
\Delta X_t &= K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, \\
r_t &= \rho_{0X} + \rho_{1X}' X_t, \\
\Delta d_t &= \delta_{0X} + \delta_{1X}' X_t,
\end{aligned} \tag{1.20}$$

where $\Sigma_X\Sigma_X'$ is the constant conditional covariance matrix of X_t, $\epsilon_t^{\mathbb{P}}, \epsilon_t^{\mathbb{Q}} \sim N(0, I_N)$, r_t is the short rate, Δd_t is the dividend growth.

Next, as this model belongs to the class of Gaussian dynamic term structure models, it is observationally equivalent up to an affine transformation of the state vector. Hence, using this feature, we can derive a canonical form of the above general co-pricing model, which is maximally flexible in the parameterization of both its $\mathbb{P}$ and $\mathbb{Q}$ distributions of X_t such that the model is identifiable. Here, no assumptions about the processes of X_t are made; only normalizations are used to ensure econometric identification. Proposition 1 shows the canonical form, and the proof is given in Appendix A1.2.

Proposition 1. Every canonical affine term structure model is observationally equivalent to the following canonical model:

$$\begin{aligned} \Delta X_t &= K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}}, \\ \Delta X_t &= K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, \\ r_t &= r_\infty^{\mathbb{Q}} + \iota' X_t, \\ \Delta d_t &= \delta_{0X} + \delta_{1X}' X_t, \end{aligned} \tag{1.21}$$

where $\epsilon_t^{\mathbb{P}}, \epsilon_t^{\mathbb{Q}} \sim N(0, I_N)$, $K_{0X}^{\mathbb{Q}} = 0$, $K_{1X}^{\mathbb{Q}}$ is in ordered real Jordan form[1] such that the diagonal elements are represented by $\lambda^{\mathbb{Q}} = (\lambda_i^{\mathbb{Q}})$s with decreasing magnitude, Σ_X is the lower triangular Cholesky decomposition of $\Sigma_X\Sigma_X'$, ι is a vector of ones.

The result of this proposition will be used in the estimation to make sure the model is not unidentified or has any over-identifying restrictions.

[1] Alternatively, $K_{1X}^{\mathbb{Q}}$ can be specified to be a diagonal matrix with real and distinct eigenvalues on the diagonal and the eigenvalues are ordered in decreasing magnitude.

1.4.2 Estimation strategy

One way to estimate the model is to maintain the assumption that the underlying state vector is completely latent as in (1.21). However, one problem with this estimation strategy is that the number of parameters that need to be estimated can potentially be very large.

An alternative estimation strategy is to estimate $\mathbb{P}$ parameters and $\mathbb{Q}$ parameters separately, as in Joslin, Singleton and Zhu (2011). It is shown below that the canonical model in terms of X_t is observationally equivalent to a unique affine model whose pricing factors $\mathcal{P}_t$ include the short rate, the first principal component of the set of equity yields that can be viewed as a portfolio of equity yields, and the dividend growth. Under the assumption that the portfolio of yields is observed without error, the estimation of the canonical model in terms of $\mathcal{P}_t$ can be carried out in a two-step procedure. First, the $\mathbb{P}$ dynamics of the state process can be estimated by an unrestricted vector autoregression (VAR). Then, taking the estimated parameters from the VAR as given, the rest of the parameters are then estimated using maximum-likelihood estimation (MLE). The advantage of this estimation procedure is that the number of parameters to be estimated is greatly reduced. Therefore, the latter strategy will be adopted in this paper.

Specifically, I assume there are three latent factors driving the economy. Applying rotation to the latent state vector, it can be shown that the canonical model in terms of X_t defined in (1.21) is observationally equivalent to a canonical model defined in terms of $\mathcal{P}_t$, which consists of the short rate, the first principal component of equity yields (portfolio of yields)

$$\mathrm{PC}_t = W y_t \tag{1.22}$$

and dividend growth, i.e.

$$\mathcal{P}_t = \begin{bmatrix} r_t \\ \mathrm{PC}_t \\ \Delta d_t \end{bmatrix}.$$

Given

$$y_t = A_X + B_X' X_t, \tag{1.23}$$

we have

$$\mathcal{P}_t = \begin{bmatrix} r_t \\ \mathrm{PC}_t \\ \Delta d_t \end{bmatrix} = \begin{bmatrix} \rho_{0X} \\ W A_X \\ \delta_{0X} \end{bmatrix} + \begin{bmatrix} \rho_{1X}' \\ W B_X' \\ \delta_{1X}' \end{bmatrix} X_t = A + B X_t. \tag{1.24}$$

Therefore, given the affine relationship between $\mathcal{P}_t$ and X_t, we can use the generic feature of Gaussian affine term structure models to find a model in terms of $\mathcal{P}_t$, which is observationally equivalent to the canonical form of X_t. This is summarized in Proposition 2. The proof can be found in Appendix A1.3.

Proposition 2. Any canonical affine term structure model as defined in (1.21) is observationally equivalent to a unique affine co-pricing model whose pricing factors $\mathcal{P}_t$ include the short rate, the portfolios of yields Wy_t and the dividend growth. Moreover, the $\mathbb{Q}$ distribution of $\mathcal{P}_t$ is uniquely determined by ($\lambda^{\mathbb{Q}}$, $r_\infty^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}), where $\lambda^{\mathbb{Q}} = (\lambda_i^{\mathbb{Q}})$s are ordered in decreasing magnitude. That is,

$$\begin{aligned}
\Delta\mathcal{P}_t &= K^{\mathbb{P}}_{0\mathcal{P}} + K^{\mathbb{P}}_{1\mathcal{P}}\mathcal{P}_{t-1} + \Sigma_{\mathcal{P}}\epsilon^{\mathbb{P}}_t, \\
\Delta\mathcal{P}_t &= K^{\mathbb{Q}}_{0\mathcal{P}}+, K^{\mathbb{Q}}_{1\mathcal{P}}\mathcal{P}_{t-1} + \Sigma_{\mathcal{P}}\epsilon^{\mathbb{Q}}_t \\
r_t &= \rho_{0\mathcal{P}} + \rho'_{1\mathcal{P}}\mathcal{P}_t, \\
\Delta d_t &= \delta_{0\mathcal{P}} + \delta'_{1\mathcal{P}}\mathcal{P}_t
\end{aligned} \tag{1.25}$$

is a canonical Gaussian affine term structure model, where $K^{\mathbb{P}}_{0\mathcal{P}}$, $K^{\mathbb{P}}_{1\mathcal{P}}$, $\rho_{0\mathcal{P}}$, $\rho_{1\mathcal{P}}$, $\delta_{0\mathcal{P}}$, $\delta_{1\mathcal{P}}$ are explicit functions of ($\lambda^{\mathbb{Q}}$, $r^{\mathbb{Q}}_{\infty}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}). The canonical form is parameterized by ($\lambda^{\mathbb{Q}}$, $r^{\mathbb{Q}}_{\infty}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}, $K^{\mathbb{P}}_{0\mathcal{P}}$, $K^{\mathbb{P}}_{1\mathcal{P}}$). And the relationship between the parameters in the two canonical forms can be shown as follows:

$$\begin{aligned}
K^{\mathbb{Q}}_{1\mathcal{P}} &= BJ(\lambda^{\mathbb{Q}})B^{-1}, \\
K^{\mathbb{Q}}_{0\mathcal{P}} &= -K^{\mathbb{Q}}_{1\mathcal{P}}A, \\
\rho_{1\mathcal{P}} &= (B^{-1})'\iota, \\
\rho_{0\mathcal{P}} &= r^{\mathbb{Q}}_{\infty} - A'\rho_{1\mathcal{P}}, \\
\delta_{1\mathcal{P}} &= (B^{-1})'\delta_{1X}, \\
\delta_{0\mathcal{P}} &= \delta_{0X} - A'\delta_{1\mathcal{P}}.
\end{aligned} \tag{1.26}$$

Next given the observational equivalence between the canonical form in X_t and the model in $\mathcal{P}_t$ outlined above, I estimate the model in $\mathcal{P}_t$ rather than in X_t. The estimation is implemented in three steps.

In Step 1, I assume there is one portfolio of yields that is measured without error, i.e. $Wy_t = Wy^{o}_t$. Here, y^{o}_t denotes the observed yields. Hence Wy_t can be measured by the first principal component of yields.

In Step 2, as the state vector

$$\mathcal{P}_t = \begin{bmatrix} r_t \\ \mathrm{PC}_t \\ \Delta d_t \end{bmatrix}$$

is observed, it can be used to estimate the $\mathbb{P}$ process of $\mathcal{P}_t$. More specifically, as $\mathcal{P}_t$ follows a VAR process

$$\Delta\mathcal{P}_t = K_{0\mathcal{P}}^{\mathbb{P}} + K_{1\mathcal{P}}^{\mathbb{P}}\mathcal{P}_{t-1} + \Sigma_{\mathcal{P}}\epsilon_t^{\mathbb{P}},$$

$K_{0\mathcal{P}}^{\mathbb{P}}$ and $K_{1\mathcal{P}}^{\mathbb{P}}$ can be estimated from an unrestricted VAR using ordinary least squares (OLS). And these OLS estimates can be viewed as MLE estimates because of the inherent separation between the parameters of the $\mathbb{P}$ and $\mathbb{Q}$ dynamics of $\mathcal{P}_t$, that is

$$f(y_t^{\mathrm{o}}|y_{t-1}^{\mathrm{o}};\Theta) = f(y_t^{\mathrm{o}}|\mathcal{P}_t;\lambda^{\mathbb{Q}}, r_\infty^{\mathbb{Q}}, \Sigma_{\mathcal{P}}, \delta_{0X}, \delta_{1X}, P^{\theta_m}) \times f(\mathcal{P}_t|\mathcal{P}_{t-1}; K_{1\mathcal{P}}^{\mathbb{P}}, K_{0\mathcal{P}}^{\mathbb{P}}, \Sigma_{\mathcal{P}}), \quad (1.27)$$

where P^{θ_m} is the conditional distribution of the measurement errors $y_t^{\mathrm{o}} - y_t$.

In Step 3, taking the estimated parameters from OLS as given, we estimate the rest of the parameters using MLE.

The detailed justification for the estimation strategy is as follows.

By Proposition 2, we can, without loss of generality, use

$$\mathcal{P}_t = \begin{bmatrix} r_t \\ \mathrm{PC}_t \\ \Delta d_t \end{bmatrix} \in \mathbb{R}^N$$

as observed factors. Suppose that the individual bond yields, y_t, are to be used in estimation and that their associated measurement errors, $y_t^o - y_t$, have the conditional distribution P^{θ_m}, for some $\theta_m \in \Theta_m$. It only requires that, for any P^{θ_m}, these errors are conditionally independent of lagged values of the measurement errors and satisfy the consistency condition

$$\mathbb{P}\left(\begin{bmatrix} r_t \\ W y_t^o \\ \Delta d_t \end{bmatrix} = \begin{bmatrix} r_t \\ \mathrm{PC}_t \\ \Delta d_t \end{bmatrix} |\mathcal{P}_t \right) = 1.$$

Then the conditional likelihood function (under $\mathbb{P}$) of the observed data (y_t^o) can be decomposed as the product of two conditional likelihood functions. The first likelihood function describes the conditional distribution of the observed yields that are measured with errors, that is, dependent on the parameters relevant for pricing ($\lambda^{\mathbb{Q}}$, $r_\infty^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}) and the distribution assumption of the measurement errors. The second likelihood function describes the conditional distribution of $\mathcal{P}_t$, which depends only on ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$, $\Sigma_{\mathcal{P}}$). We have

$$f(y_t^o | y_{t-1}^o; \Theta) = f(y_t^o | \mathcal{P}_t; \lambda^{\mathbb{Q}}, r_\infty^{\mathbb{Q}}, \Sigma_{\mathcal{P}}, \delta_{0X}, \delta_{1X}, P^{\theta_m}) \\ \times f(\mathcal{P}_t | \mathcal{P}_{t-1}; K_{1\mathcal{P}}^{\mathbb{P}}, K_{0\mathcal{P}}^{\mathbb{P}}, \Sigma_{\mathcal{P}}).$$

Now if we assume $\mathcal{P}_t$ is conditionally Gaussian, then the conditional $\mathbb{P}$ likelihood of $\mathcal{P}$, i.e. the second part of (1.28) can be expressed as

$$f(\mathcal{P}_t | \mathcal{P}_{t-1}; K_{1\mathcal{P}}^{\mathbb{P}}, K_{0\mathcal{P}}^{\mathbb{P}}, \Sigma_{\mathcal{P}}) \\ = (2\pi)^{-N/2} \left|\Sigma_{\mathcal{P}}\right|^{-1} \times \exp(-\tfrac{1}{2} \left\| \Sigma_{\mathcal{P}}^{-1} (\mathcal{P}_t - E_t[\mathcal{P}_t]) \right\|^2), \quad (1.28)$$

where $E_{t-1}[\mathcal{P}_t] = K_{0\mathcal{P}}^{\mathbb{P}} + (I + K_{1\mathcal{P}}^{\mathbb{P}})\mathcal{P}_{t-1}$.

Moreover, Zellner (1962) shows that, conditional on $t = 0$ information, the parameters $K_{1\mathcal{P}}^{\mathbb{P}}$ and $K_{0\mathcal{P}}^{\mathbb{P}}$ that maximize this likelihood function are their OLS estimates, i.e.

$$\begin{aligned}(K_{1\mathcal{P}}^{\mathbb{P}}, K_{0\mathcal{P}}^{\mathbb{P}}) &= \operatorname{argmax} \sum_{t=1}^{T} f(y_t^{o} | y_{t-1}^{o}; K_{1\mathcal{P}}^{\mathbb{P}}, K_{0\mathcal{P}}^{\mathbb{P}}, \Sigma_{\mathcal{P}}) \\ &= \operatorname{argmin} \sum_{t=1}^{T} \left\| \Sigma_{\mathcal{P}}^{-1} (\mathcal{P}_t - E_t[\mathcal{P}_t]) \right\|^2 . \end{aligned} \tag{1.29}$$

Hence, ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$) can be estimated from time series of $\mathcal{P}_t$ alone, and their OLS estimates are globally optimal, i.e. they are equal to their maximum likelihood estimates. Hence ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$) will no longer need to be estimated using MLE, and the canonical form in $\mathcal{P}$ is now parameterized by ($\lambda^{\mathbb{Q}}$, $r_{\infty}^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}) rather than by ($\lambda^{\mathbb{Q}}$, $r_{\infty}^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}, $K_{0\mathcal{P}}^{\mathbb{P}}$, $K_{1\mathcal{P}}^{\mathbb{P}}$), effectively eliminating the dependence on ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$). The separation is formally shown in Proposition 3.

Proposition 3. Using the observed factor, $\mathcal{P}_t = \mathcal{P}_t^{o} \in \mathbb{R}^N$, the maximum likelihood estimates of ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$) are given by their OLS estimates. Moreover, the canonical form of $\mathcal{P}_t$ in Proposition 2 is now parameterized by ($\lambda^{\mathbb{Q}}$, $r_{\infty}^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}), effectively eliminating the dependence on ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$).

Moreover, the sample estimates of $\Sigma_{\mathcal{P}}$ can be used as starting values for their MLE estimation, reducing the estimation time of these parameters. The reduction in the number of parameters need to be estimated using MLE increases with the number of factors that are assumed to drive the economy. For example, with an N-factor model, we no longer need to estimate the $N(N+1)$ parameters that come from ($K_{1\mathcal{P}}^{\mathbb{P}}$, $K_{0\mathcal{P}}^{\mathbb{P}}$), which greatly reduces the estimation time.

1.4.3 Data

This subsection describes the data used in this paper, especially the three observed factors assumed to drive the whole economy, i.e. the short rate, dividend growth and portfolio of equity yields. The procedure used to remove seasonality in the data and the way the equity portfolio is constructed are also discussed.

Short rate

First, the short rate is included as a state vector in the VAR under $\mathbb{P}$. Figure 1.1 plots the time series of the one-month interest rate provided by van Binsbergen, Brandt and Koijen (2012) for the sample period. Generally, interest rates exhibit extremely high persistence. The first-order autocorrelation coefficient is usually close to one and a unit root test could not reject the existence of a stochastic trend (see Goodfriend (1991) and Jardet, Monfort and Pegoraro (2013)). However, in this paper, I follow the majority of the literature (see Bauer, Rudebusch and Wu (2012)) and assume that the short rate is stationary. This is because if we were to allow the short rate to be nonstationary, a unit root will mean that the short rate will never revert back to its long-run mean and an explosive root will mean that the short rate will inevitably drop below the zero lower bound; both cases contradict the fact that nominal interest rates are bounded above zero and remain within a certain range.

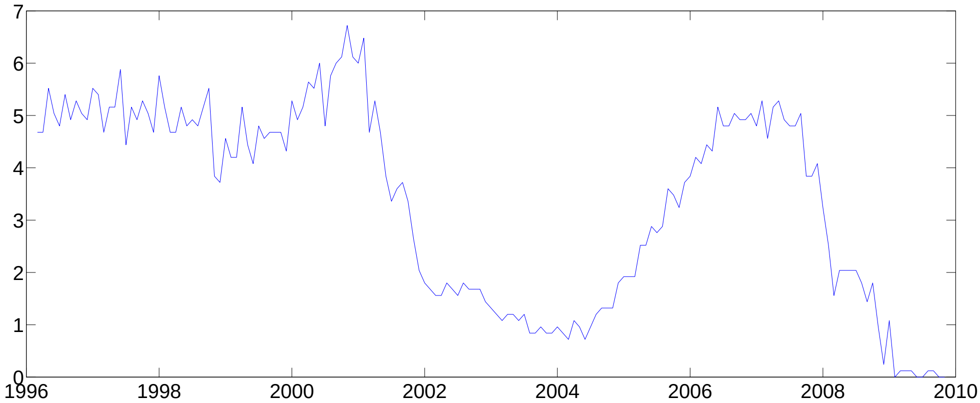


Figure 1.1: Short rate, in annual percentages, 1996m2:2009m10

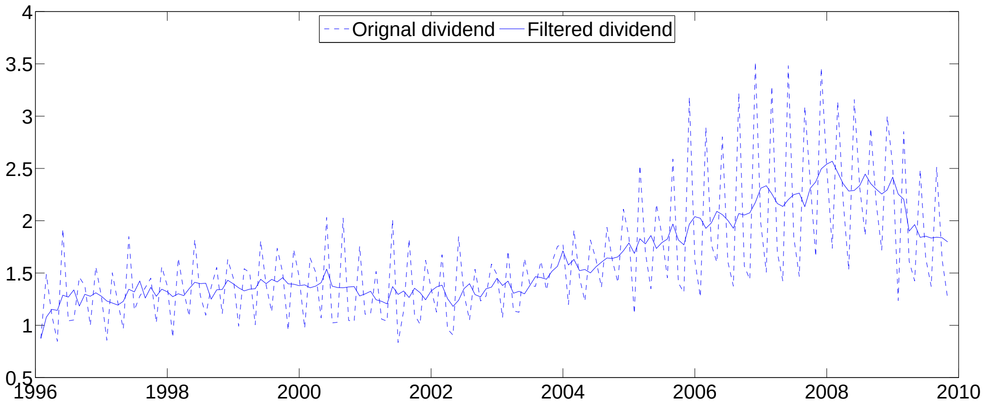
Dividend and dividend growth

Monthly dividends from January 1996 to October 2009 are provided by van Binsbergen, Brandt and Koijen (2012). The dashed line in the top panel of Figure 1.2 shows monthly dividends and the dashed line in the middle panel shows the monthly log dividend growth rates. We can see that the level of dividend is nonstationary but log dividend growth rates are stationary. However, monthly dividends exhibit seasonality in both level and log growth rates, due to the fact that most companies pay dividends on a quarterly basis.

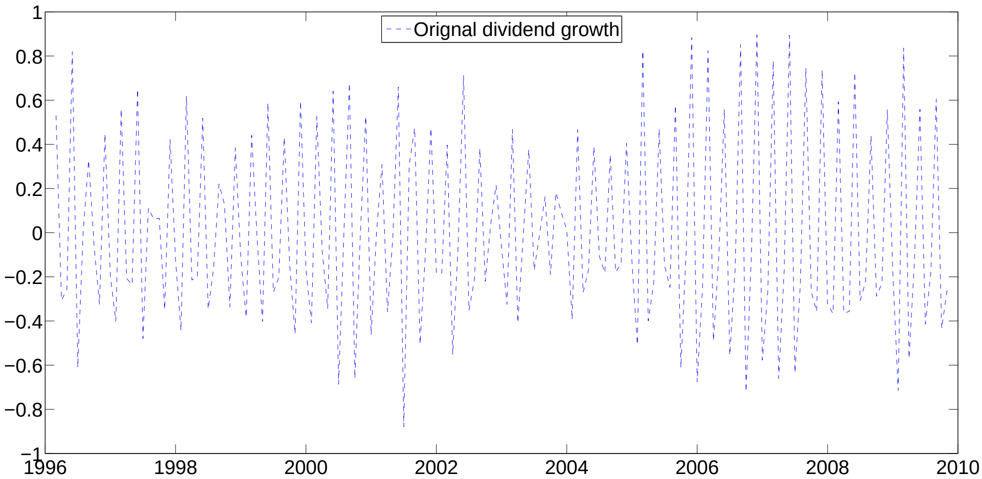
If we were to use original monthly dividend in the estimation, because of its own seasonality, not only will it cause dividend growth to be seasonal, it will also cause equity yields to be seasonal, which will be seen later in this section. And the seasonality will cause a problem for the estimation. This is because seasonality means shocks are not independent and identically distributed (i.i.d.), which will violate the standard i.i.d. assumption in estimation. In addition, as seasonality is not explicitly modeled in the paper, using the data with seasonality in the estimation will bias the estimates of the model. Therefore, we must filter out the seasonality in the data before it enters into the estimation. Moreover, filtering out seasonality can also help us see more clearly the underlying trend of the data.

To remove the seasonality in dividend, I deseasonalize the dividend itself, then use the deseasonalized dividends to calculate dividend growth, and to scale the prices of zero-coupon equities to calculate equity yields.

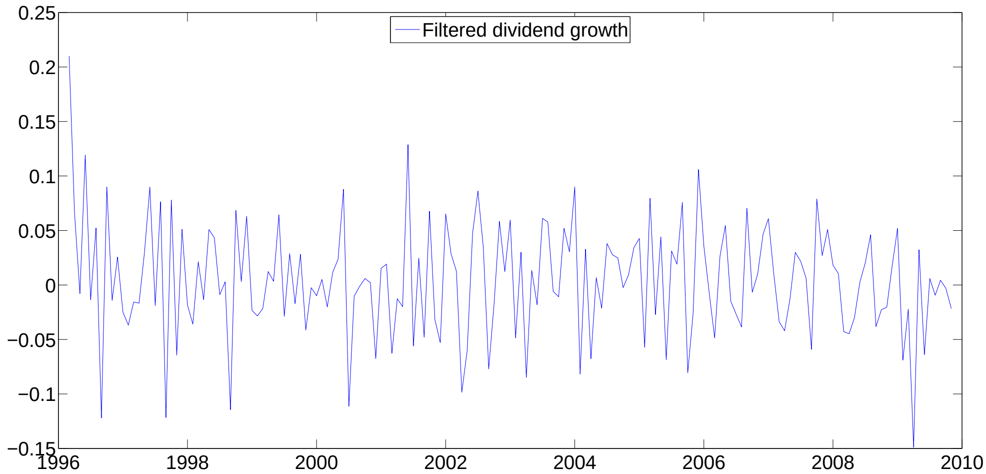
More specifically, I choose a simple moving average over three months to filter out the seasonality in the monthly dividend. The solid line in the top panel of Figure 1.2 shows the filtered monthly dividends. We can see that, using the above method of filtering, we are able to remove the quarterly seasonality that was previously present in the monthly dividends, and we can also see the trend of the monthly dividends more clearly. The solid line in the bottom panel of Figure 1.2 shows the monthly log dividend growth rates based on the deseasonalized monthly dividends. By using deseasonalized dividends, the quarterly seasonality that was previously affecting dividend growth is also reduced.



(a) Monthly dividends, D_t, original and filtered, 1996m1:2009m10



(b) Monthly log dividend growth rates, $\Delta d_t = \log(D_t/D_{t-1})$, original, 1996m2:2009m10



(c) Monthly log dividend growth rates, $\Delta d_t = \log(D_t/D_{t-1})$, filtered, 1996m2:2009m10

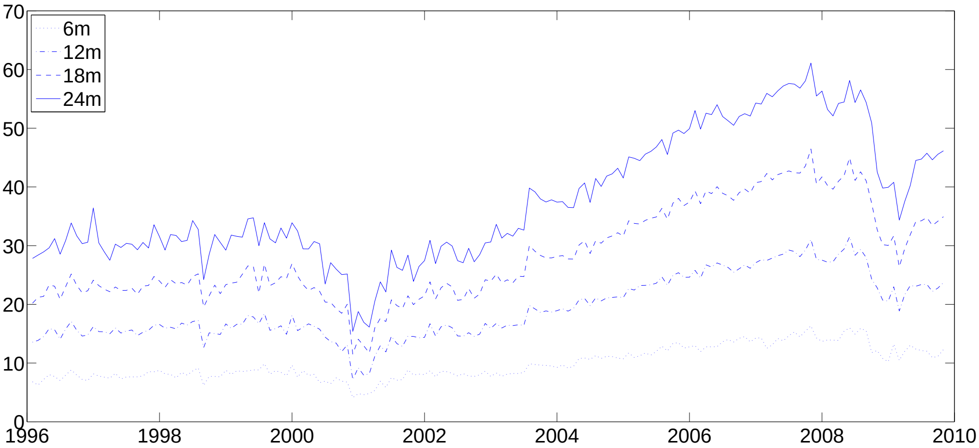
Figure 1.2: Monthly dividends and log dividend growth rates, original and filtered.

Equity yield portfolio

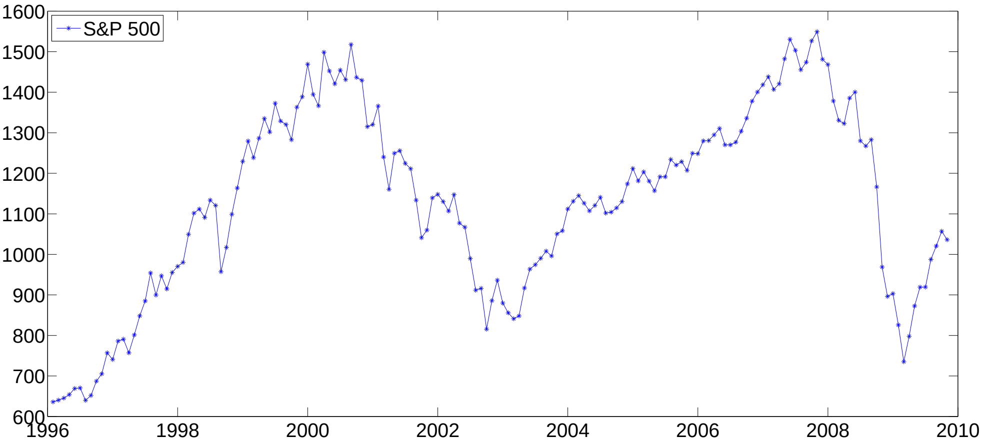
Finally, an "equity yield portfolio" is included as a state factor in the estimation. The factor is the first principal component of equity yields, which is a linear combination of equity yields and hence can be viewed as a portfolio of equity yields. Equity yields in this paper are defined by equation (1.13), i.e.

$$y^{\mathrm{d}}_{nt} = -\frac{1}{n} \ln \frac{P^{\mathrm{d}}_{nt}}{D_t}.$$

P^{d}_{nt} is the price of zero-coupon equities and the data on zero-coupon equities are calculated from the prices of dividend strips provided by van Binsbergen, Brandt and Koijen (2012). van Binsbergen, Brandt and Koijen match the S&P 500's put and call prices of the same maturity and use put–call parity to back out strip prices. Specifically, the S&P 500 index is decomposed into a portfolio of dividend strips, which entitles the holder to the realized dividends of the index with maturities of six, twelve, eighteen and twenty-four months. For example, from the dataset we know that in January 1996 the price of the dividend strip with twelve-month maturity is $13.56. Buying this dividend strip in January 1996 entitles the buyer to a dividend paid out at each month end from February 1996 to January 1997. And by January 1997 the total amount of dividends paid out in the past 12 months reaches $14.97. The data for the dividend strips are available monthly from January 1996 to October 2009. The times series of the prices of dividend strips and S&P 500 index are plotted in the top and bottom panels of Figure 1.3. From the figure, we can see that the prices of dividend strips are monotonically increasing with maturity, as violations of this would lead to arbitrage opportunities. Ideally, to extract the prices of zero-coupon equities, we need prices of strips of two adjacent maturities. For example, to obtain the price of monthly zero-coupon equity with a maturity of twenty-four months, we need the prices of the dividend strips for the twenty-four-month and twenty-three-month maturities. However, strips of only four maturities, each six months apart, are available. This means that, to calculate the prices of monthly zero-coupon equities, we need to interpolate the prices of the dividend strips in between using the observed relationship between the prices of dividend strips and their maturities. The method in this paper is to use nonlinear curve fitting to fit six observed strip prices with maturities of 0, 6, 12, 18, 24 months and ∞; the last strip price is actually the S&P 500 stock index. This is because in theory the S&P 500 can be viewed as the sum of zero-coupon equity prices from maturity one to maturity infinity. Hence it can be viewed as a strip price of maturity at infinity. Since the prices of the dividend strip with maturity at zero months and infinity months are zero and the S&P 500 stock index, the nonlinear curve is constructed to always satisfy three properties. First, it has to go through the origin. Second, it has to be monotonically increasing to rule out any arbitrage opportunities. Finally, the curve should asymptotically approach the S&P 500 index as the maturity goes to infinity.



(a) Prices of dividend strips

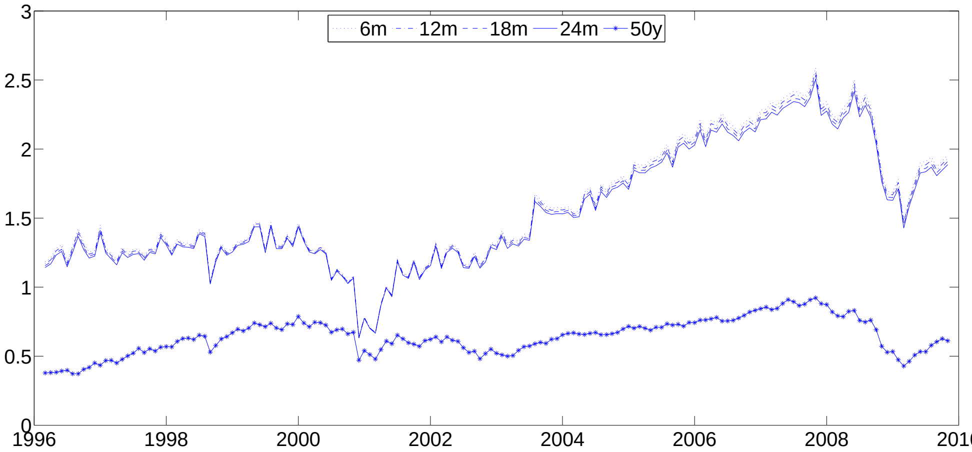


(b) Level of the S&P 500

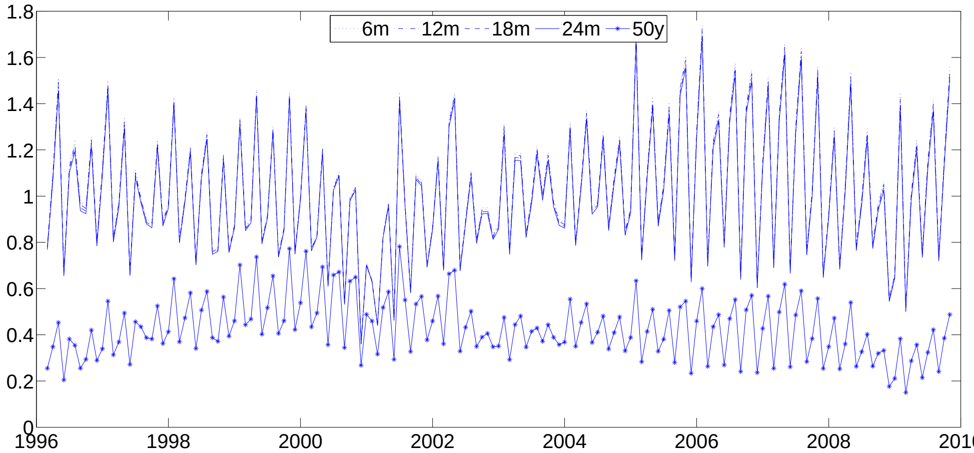
Figure 1.3: Prices of dividend strips and the level of the S&P 500, 1996m1:2009m10

Ideally, for the MLE estimation, we want to have yields across a range of maturities to pin down the equity term structure. When yields are not available, we could choose to match prices that are cumulative exponentials of the yields. Given that we have a fitted curve that matches the strip prices and the index very well, the ideal would be to match the strip prices and the market index. However, incorporating the index into the estimation is technically difficult. This is because matching the market index requires matching the sum of all the model-generated zero-coupon equity prices from maturity one to maturity infinity with the observed index. The assumption of which period we take as infinity is arbitrary but will nevertheless be large for the market index. Hence, in practice, matching the market index under the affine model will be too computationally intensive. Alternatively, given the fact that the entire strip price curve is generated, we can extract yields directly from the curve. We could use the fitted curve to extract information regarding the long end of the equity term structure. More specifically, a zero-coupon equity of a long-maturity (fifty-year) yield could be taken from the estimated nonlinear curve and be used as the data on the long-maturity yield in the estimation. This equity yield at long maturity will not be a perfect substitute for the market index, but it will nevertheless contain information on the long end of the equity term structure. By having the equity yield at long maturity, the estimation will have a greater ability to fit the observed market index.

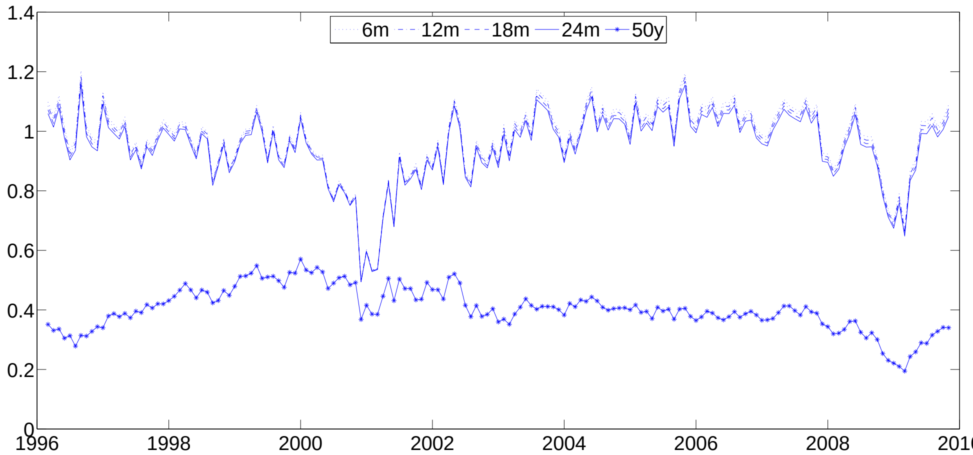
By subtracting the fitted strip prices of two adjacent maturities, we obtain the prices of the zero-coupon equities. The prices of zero-coupon equities (6m, 12m, 18m, 24m and 50y) from January 1996 to October 2009 are plotted in the top panel of Figure 1.4. We can see that the prices of zero-coupon equities follow the same trend as those of dividend strips. And because the four zero-coupon equities' maturities are close to each other, their prices are generally at the same level. As pointed out by van Binsbergen, Brandt and Koijen (2012) (BBK), dividend strip prices are nonstationary over time; they scale the four dividend strip prices by the level of the S&P 500 index to obtain stationary data series. Indeed, the prices of the zero-coupon equities are nonstationary, as seen from the top panel of Figure 1.4. However, if we scale zero-coupon equity prices by the dividend provided by BBK, then P^{d}_{nt}/D_t is stationary, as shown by the middle panel of Figure 1.4. The prices of zero-coupon equities scaled by monthly dividends provided in BBK are stationary but seasonal, which is caused by the seasonality in monthly dividends as shown in the previous section. The bottom panel of Figure 1.4 shows that zero-coupon equity prices scaled by deseasonalized dividends are stationary and are no longer affected by seasonality. We can then use these to extract yields.



(a) Prices of zero-coupon equities



(b) Prices of zero-coupon equities scaled by original dividend



(c) Prices of zero-coupon equities scaled by filtered dividend

Figure 1.4: Prices of zero-coupon equities, 1996m1:2009m10

Next, we extract the principal component from the group of equity yields of maturities 6m, 12m 18m, 24m and a long maturity of 50 years, and the first principal component explains over 99% of the total variations in this group of equity yields. Figure 1.5 plots the time series of the five yields. The two spikes in 2001 and 2009 are indicative of two economic recessions. Figure 1.6 plots the time series of the first principal component of the equity yields.

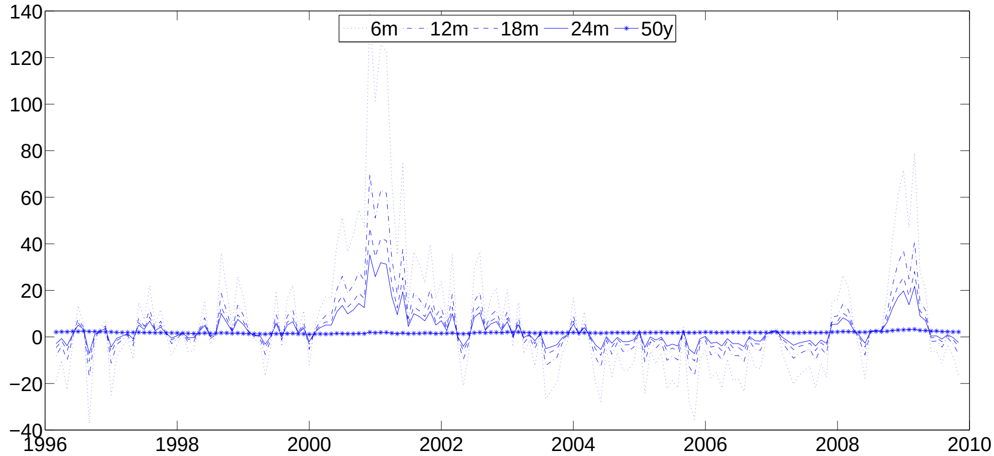


Figure 1.5: Annual percentage yields of zero-coupon equities

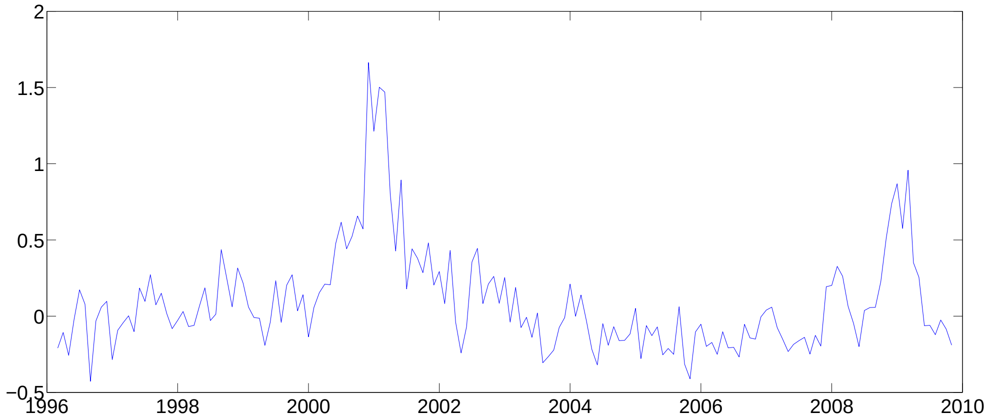


Figure 1.6: First principal component of zero-coupon equities

1.4.4 Estimation results

Model parameters

As the likelihood function is optimized over $(\lambda^{\mathbb{Q}}, r^{\mathbb{Q}}_{\infty}, \Sigma_{\mathcal{P}}, \delta_{0X}, \delta_{1X})$, these parameters' estimates are shown in Table 1.1 and Table 1.2. Estimates are in annual numbers with their standard deviations in parentheses. From the estimates of the $(\lambda_i^{\mathbb{Q}}+1)$s, we can see that the estimated state process has real and distinct eigenvalues, and because all $(\hat{\lambda}_i^{\mathbb{Q}}+1)$s are between zero and one, the estimated state process is stationary under $\mathbb{Q}$. The initial value of $\Sigma_{\mathcal{P}}$ is obtained from the VAR estimate of the $\mathbb{P}$ process of the state vector, and the final estimate of $\Sigma_{\mathcal{P}}$ obtained from the MLE is reported in Table 1.2. Table 1.3 reports the OLS estimates of $K_{0\mathcal{P}}^{\mathbb{P}}$ and $K_{1\mathcal{P}}^{\mathbb{P}}+I$.

Table 1.1: Maximum likelihood estimates of the risk-neutral parameters, equity term structure

	$\lambda_1^{\mathbb{Q}}+1$	$\lambda_2^{\mathbb{Q}}+1$	$\lambda_3^{\mathbb{Q}}+1$	$r_\infty^{\mathbb{Q}}$
Estimate	0.9967	0.3761	0.3068	−670.2887
Standard deviation	(0.0010)	(0.0170)	(0.6789)	(0.3190)
	δ_{0X}	$\delta_{1X,1}$	$\delta_{1X,2}$	$\delta_{1X,3}$
Estimate	−6.7241	0.9983	−0.7104	−1.3359
Standard deviation	(0.0117)	(0.0030)	(0.8033)	(1.1760)

Table 1.2: Maximum likelihood estimates of the conditional covariance, equity term structure

	$\sum_{\mathcal{P},11}$		
Estimate	0.4655		
Standard deviation	(0.0304)		
	$\sum_{\mathcal{P},21}$	$\sum_{\mathcal{P},22}$	
Estimate	−0.6509	19.6996	
Standard deviation	(1.2707)	(2.0452)	
	$\sum_{\mathcal{P},31}$	$\sum_{\mathcal{P},32}$	$\sum_{\mathcal{P},33}$
Estimate	5.0893	23.7969	51.0953
Standard deviation	(4.5313)	(7.0941)	(3.6709)

Table 1.3: Maximum likelihood estimates of the physical parameters, equity term structure

	$K^{\mathbb{P}}_{0\mathcal{P},1}$	$K^{\mathbb{P}}_{0\mathcal{P},2}$	$K^{\mathbb{P}}_{0\mathcal{P},3}$
Estimate	0.0585	−2.9820	0.9645
Standard deviation	(0.0764)	(3.4513)	(9.0396)
	$K^{\mathbb{P}}_{1\mathcal{P},11}+1$	$K^{\mathbb{P}}_{1\mathcal{P},12}$	$K^{\mathbb{P}}_{1\mathcal{P},13}$
Estimate	0.9820	−0.0023	−0.0008
Standard deviation	(0.0200)	(0.0011)	(0.0006)
	$K^{\mathbb{P}}_{1\mathcal{P},21}$	$K^{\mathbb{P}}_{1\mathcal{P},22}+1$	$K^{\mathbb{P}}_{1\mathcal{P},23}$
Estimate	1.6157	0.7616	−0.0425
Standard deviation	(0.9032)	(0.0491)	(0.0263)
	$K^{\mathbb{P}}_{1\mathcal{P},31}$	$K^{\mathbb{P}}_{1\mathcal{P},32}$	$K^{\mathbb{P}}_{1\mathcal{P},33}+1$
Estimate	2.6735	−0.5392	−0.2203
Standard deviation	(2.3657)	(0.1286)	(0.0690)

Equity yields

Given the model parameters, we can generate the model-implied equity yields and compare them with the equity yields observed in the data. The top panel of Table 1.4 reports the summary statistics for equity yields in the data. We can see that the mean of equity yield is decreasing in maturity and the standard deviation of equity yield is also decreasing in maturity. The bottom panel of Table 1.4 reports the summary statistics for equity yields implied by the model. The estimated yields can match the observed yields very well in terms of both the first moment and the second moment. Hence these results demonstrate that this estimation framework, originally designed to match bond yields, is able to match the basic features of equity yields as well.

Table 1.4: Estimation results: moments comparison for equity yields, annual percentages, equity term structure

Maturity (years)	0.5	1	1.5	2	long maturity (50)
Panel A: Data					
Mean	6.91	4.38	3.53	3.11	1.85
Standard deviation	28.60	14.17	9.36	6.96	0.38
Panel B: Model					
Mean	6.91	4.38	3.53	3.11	1.85
Standard deviation	28.60	14.21	9.34	6.90	0.10

Return of the dividend strip and the aggregate market

BBK construct the return of the short-term strip and compare it with the return of the market. They find that the return of the short-term strip is much higher than that of the market. Given the return of the market can be viewed as a weighted average of the return of the short-term strip and the return of the long-term strip, the above observation implies that the return of the short-term strip is higher than the return of the long-term strip, i.e. the equity term structure is downward sloping. This can be illustrated by Table 1.5.

Table 1.5: Data: summary statistics of the return of the short-term strip and the market, annual percentages

	$R_{\text{short strip}}$	$R_{\text{short strip}} - R_f$	R_{market}	$R_{\text{market}} - R_f$
Mean	13.90	10.53	6.67	3.29
Standard deviation	27.03	27.04	16.26	16.22
Sharpe ratio	0.39	—	0.20	—

We can see that the annual mean return of the short-term strip is 13.90% whereas the annual mean return of the market is only 6.67%. The annual mean excess return of the short-term asset is 10.53% whereas the annual mean excess return of the market is only 3.29%. All these imply a downward sloping equity term structure. In addition, the Sharpe ratio of the short-term asset is higher than the Sharpe ratio of the market despite its high volatility.

We can also check whether this evidence can be generated by the model in the paper. In particular, the two returns are calculated as follows. For the short-term strip, its return can be calculated as

$$R^{\mathrm{s}}_{t+1} = \frac{P^{\mathrm{s}}_{n-1,t+1} + D_{t+1}}{P^{\mathrm{s}}_{n,t}} - 1, \tag{1.30}$$

where $P^{\mathrm{s}}_{n,t}$ is the time-t price of the dividend strip of maturity n, and $P^{\mathrm{s}}_{n,t} = \sum_{i=1}^{n} P^{\mathrm{d}}_{i,t}$, that is, it is the cumulative price of zero-coupon equities of maturity 1 to maturity n. For the market return, given the time series of zero-coupon equities of maturities one to infinity, we can construct the equity index of each period, together with each period's aggregate dividend. The return on the market index can be computed using

$$R^{\mathrm{m}}_{t+1} = \frac{P^{\mathrm{m}}_{t+1} + D_{t+1}}{P^{\mathrm{m}}_{t}} - 1. \tag{1.31}$$

Using the two formulas above, Table 1.6 shows that the above features within the BBK data can be matched well by our model. In BBK's data, the maturities of the short-term strip vary between 1.3 years and 1.9 years. Here, the returns of the strip of maturity 15 months to 23 months are listed in the table to match the maturities used by BBK. The means and the standard deviations of the maturities are largely in line with those provided in BBK. The last column of the table provides the model-generated market return, which closely matches that in the data. Table 1.7 also lists the excess return, standard deviation and the Sharpe ratio of the short-term strip and the market. All can be shown to match the data well.

Table 1.6: Summary statistics of the return of the short-term strip and the market, equity term structure, annual percentages

	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	11.79	11.66	11.55	11.44	11.35	11.26	11.19	11.11	11.05	6.59
Standard deviation	26.15	26.26	26.35	26.43	26.51	26.57	26.62	26.67	26.71	16.17
Sharpe ratio	0.32	0.32	0.31	0.31	0.30	0.30	0.29	0.29	0.29	0.20

Table 1.7: Summary statistics of the excess return of the short-term strip and the market, equity term structure, annual percentages

	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	8.43	8.30	8.19	8.09	7.99	7.91	7.83	7.76	7.69	3.23
Standard deviation	26.17	26.28	26.37	26.45	26.52	26.59	26.64	26.69	26.73	16.18

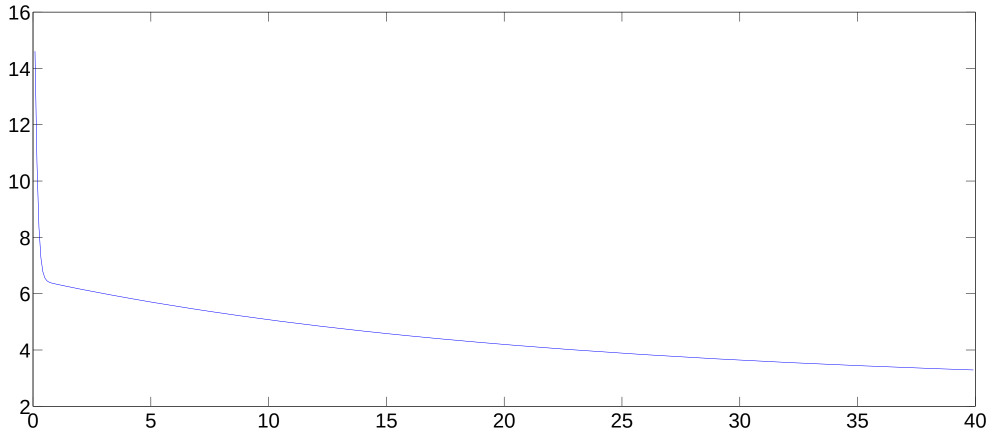
Risk premiums of zero-coupon equities

In this section, we look at the equity term structure from another perspective. In particular, we look at the risk premiums of zero-coupon equities, that is, the one-period return of these assets in excess of the risk-free rate. The reason we look at these quantities is because Lettau and Wachter (2007) use them as empirical support for the value premium. Their rationale is that if we think of value stocks as short-horizon equities since their cash flows are weighted more towards the present, and think of growth stocks as long-horizon equities since their cash flows are weighted more towards the future, then we can only observe the value premium if we see that zero-coupon equities of shorter maturities have higher premiums than zero-coupon equities with longer maturities. Value premium implies a downward sloping equity term structure.

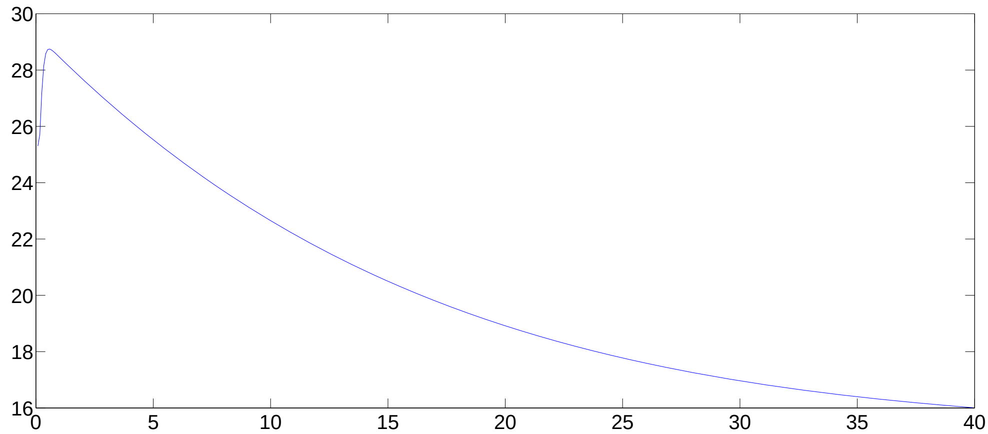
The risk premium or the one-period excess return on zero-coupon equity is calculated using the following equation:

$$R^{\mathrm{d}}_{n,t+1} - R_f = \frac{P^{\mathrm{d}}_{n-1,t+1}}{P^{\mathrm{d}}_{nt}} - R_f. \tag{1.32}$$

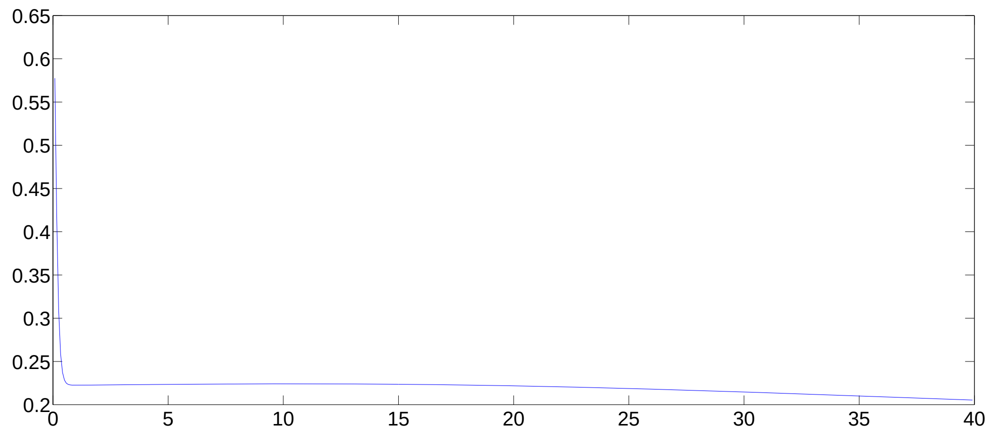
The top panel of Figure 1.7 shows the estimated average annual risk premiums for the zero-coupon equities. The middle panel shows the annual volatility and the bottom panel shows the Sharpe ratios. Zero-coupon equities with maturities up to forty years are examined (the same set of maturities as examined in Lettau and Wachter (2007)). We can see that the basic features of the model's implied excess returns of zero-coupon equities are largely consistent with the calibration of Lettau and Wachter (2007). Although the risk premiums in their calibration are generally higher than those generated from the estimation in this paper, both have risk premiums of zero-coupon equities decline with maturity. The return volatility initially increases with maturity, then decreases. The unconditional Sharpe ratio generally decreases with maturity.



(a) Estimated annual mean risk premium for ZCE, maturity in years



(b) Estimated annual volatility of risk premium for ZCE, maturity in years



(c) Estimated annual Sharpe ratios of risk premium for ZCE, maturity in years

Figure 1.7: Estimated risk premiums of zero-coupon equities, equity term structure

Therefore, we can see that, for the purpose of matching dividend strips or zero-coupon equities' yields and returns characteristics, the current model setup is sufficient.

1.5 Conclusion

This chapter examines the term structure of equities. Using the observed prices of dividend strips on the stock market index in van Binsbergen, Brandt and Koijen (2012), the prices of zero-coupon equities are extracted and their yields and returns characteristics are documented. An affine term structure model with three risk factors (the short rate, an equity yield portfolio and the aggregate dividend growth) is used to model the term structure of the equities. Identification is ensured by extending the canonical form for bonds outlined in Joslin, Singleton and Zhu (2011) to a canonical form for equities. The model is estimated, and model-implied equity yields and returns are shown to match the data well. However, as the model is estimated without taking into account data on zero-coupon bond yields, the model-implied risk-neutral long-run mean of the short rate is implausible. The second chapter takes this into account and estimates the bond and equity yield curves jointly using data on both zero-coupon bonds and zero-coupon equities.

Chapter 2

Estimating a Unified Framework of Co-Pricing Stocks and Bonds

2.1 Abstract

This chapter estimates a maximal identifiable affine term structure model that explains the joint prices of stocks and bonds. Using the test assets of Treasury bonds and dividend strips, it is shown that the estimated model can generally match the time series and cross-sectional properties of zero-coupon bonds, zero-coupon equities and the aggregate stock index. Moreover, imposing restrictions prevalent in the co-pricing literature on the maximal model enhances certain features of the model such as the high return of the short-term dividend strip, but reduces the model's ability to fit other aspects of the data such as the level of the market risk premium.

2.2 Introduction

In the literature stocks and bonds have been priced well under separate frameworks. Stocks are usually priced by using equilibrium models. Examples of such models include the external habit formation model of Campbell and Cochrane (1999), the long-run risks model of Bansal and Yaron (2004), and the rare disasters model of Gabaix (2012), which builds upon the work of Barro (2009). Bonds are more often priced by affine term structure models. Examples of such include Duffie and Kan

(1996), Dai and Singleton (2000) and Ang and Piazzesi (2003). However, if investors have access to both stocks and bonds, then the assumption of no-arbitrage will imply cross-market restrictions on the pricing kernel or the stochastic discount factor, which can be used to price all assets in the market. Hence there should exist a unified framework that is able to price both stocks and bonds.

There is now a small and growing literature that tries to use the no-arbitrage affine framework to jointly price stocks and bonds, although each paper has its own focus. In Lettau and Wachter (2011), the focus is on matching an upward sloping bond yield curve and a downward sloping equity yield curve. Koijen, Lustig and Van Nieuwerburgh (2013) is a reduced-form model that uses a cyclical factor to price the book-to-market sorted stock portfolios and maturity sorted bond portfolios. Ang and Ulrich (2012) decomposes expected equity returns into various yields and risk premiums. The key advantage that these affine models have in common is tractability: both stock yields and bond yields are affine functions of the state vector, and the loadings on the state vector are functions of the model parameters. Hence we can easily see how the state vector affects yields analytically.

However, as these papers all belong to the class of Gaussian affine term structure models, they also share a common generic feature: as shown in Dai and Singleton (2000), any affine transformations of the state process are observationally equivalent. That is, suppose we have two models, the same in all aspects except that the state process in one model is an affine transformation of the state process in the other model. Then the two models will generate exactly the same asset prices. In other words, given the same set of asset prices, there exists an infinite number of models that can generate this set of asset prices. Therefore, these models are generally not identified without imposing restrictions. To impose the minimum number of restrictions on the state process such that the model is identified, we can follow Joslin, Singleton and Zhu (2011) (JSZ) to normalize all the models to a canonical form, in which the state vector X_t is entirely latent. As a result of this normalization, given the same set of asset prices, there is only one model in the canonical form that can generate this set of asset prices. And all the state vectors in the models that can generate this set of asset prices are just rotations (affine transformations) of X_t. Therefore, if we denote the state vector in the existing papers of co-pricing stocks and bonds as Z_t, then the rotation between Z_t and X_t and the minimum number of restrictions on the process of X_t imply that the process of Z_t will also face parameter restrictions. As the existing models of co-pricing stocks and bonds are usually calibrated and fail to take into account these restrictions in the calibration, this could lead to some of the model parameters being over-restricted (in the sense that the canonical form implies these parameters should be freely estimated but they are instead restricted to zeros or ones) or not identified (in the sense that the canonical form implies these parameters should be restricted to zeros or ones but instead they are freely estimated), which may cause spurious predictions of asset pricing moments.

In this chapter, I develop and estimate a co-pricing model that jointly prices nominal bonds and equities, taking into account the restrictions implied by the canonical form. The model belongs to the class of Gaussian affine term structure models in the sense that all shocks are normally distributed and asset prices are exponential affine in the underlying state vector that drives the economy. Existing papers under this framework only use data on zero-coupon bonds and the aggregate market index, but this paper also utilizes zero-coupon equities in the estimation.

I estimate the model following Joslin, Singleton and Zhu (2011). I first normalize the general model to the canonical form. Then I show that the canonical form is observationally equivalent to another model, in which the $N \times 1$ state vector includes dividend growth and $N-1$ principal components (PCs) extracted from the yields of zero-coupon bonds and the yields of zero-coupon equities, which can be seen as portfolios of these yields. Including dividend growth in the state vector is motivated by the fact that dividend growth cannot be spanned by the PCs of yields: it can be shown that, when dividend growth is regressed on a constant and the $N-1$ PCs of yields, the R-squared is very low, i.e. variation in dividend growth cannot be explained by the PCs of yields. If we do not include dividend growth as a state factor, as the pricing function of zero-coupon equities requires dividend growth to be expressed as an affine function of the state vector we would implicitly assume dividend growth is an affine function of the PCs of yields, which is not the case in the data. Therefore, we must include dividend growth explicitly in the state vector.

The model is estimated using maximum likelihood estimation (MLE) following JSZ. Under the assumption that the PCs (portfolios of yields) are observed without error, the estimation can be carried out in a two-step procedure. First, the $\mathbb{P}$ dynamics of the state process can be estimated by an unrestricted VAR. Then, taking the estimated parameters from the VAR as given, the rest of the parameters are then estimated using MLE. Data on bond yields used in the estimation are Constant Maturity Treasury yields, and the empirical counterparts of zero-coupon equities are calculated from the dividend strips data provided in van Binsbergen, Brandt and Koijen (2012). The estimated model can match the time series and cross-sectional properties of asset pricing moments for both stocks and bonds well. Moreover, the estimation results show that it is important to take into account the above restrictions to match all the asset pricing moments. It will be seen later that imposing additional restrictions on top of the identifying restrictions implied by the maximal identifiable model would strengthen some asset pricing features. However, this is achieved at the expense of not matching the other asset pricing features.

By estimating a model of co-pricing stocks and bonds, this paper fills the gap in the literature in which bonds and stocks are usually priced separately. Moreover, in the estimation, I use a maximal identifiable model to make sure I do not impose additional restrictions on the model that may lead to spurious results. Chernov and Mueller (2012) also estimate a model for the bond market, guided by the maximal identifiable model derived from Joslin (2006). Regarding data, I use the dataset on dividend strips provided in van Binsbergen, Brandt and Koijen (2012) to empirically estimate zero-coupon equities, a concept defined as, for example, in Lettau and Wachter (2007) but data on it was missing in the literature. Bekaert and Grenadier (2001) estimate an affine model of co-pricing stocks and bonds as well, but they use only data on the aggregate market index, without using the prices of zero-coupon equities.

The rest of the chapter is structured as follows: Section 2.3 outlines the Gaussian affine term structure model that is able to price bonds and equities, and its canonical form that is maximally identifiable. The section also describes the data and the estimation strategy and shows the estimation results. Section 2.4 compares this paper with an existing paper of co-pricing stocks and bonds to illustrate the importance of taking into account the restrictions implied by the maximal identifiable model. Section 2.5 concludes.

2.3 Model and estimation

I assume the economy in this chapter is driven by a latent state vector X_t, which follows a VAR(1) process under both the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$:

$$\Delta X_t = K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}},$$
$$\Delta X_t = K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}},$$

where X_t is an $N \times 1$ vector, $K_{0X}^{\mathbb{P}}$ and $K_{0X}^{\mathbb{Q}}$ are $N \times 1$ vectors, $K_{1X}^{\mathbb{P}}$, $K_{1X}^{\mathbb{Q}}$ and Σ_X are $N \times N$ matrices and both $\epsilon_{t+1}^{\mathbb{P}}$ and $\epsilon_{t+1}^{\mathbb{Q}}$ are $N \times 1$ vectors of independent shocks with mean zero and unit variance.

I also assume the short rate and the dividend growth are driven by X_t, as follows:

$$r_t = \rho_{0X} + \rho_{1X}' X_t,$$

where ρ_{0X} is a scalar and ρ_{1X} is an $N \times 1$ vector.

The level of the aggregate dividend of the economy is denoted by D_t. Let $d_t = \log D_t$, and the log dividend growth rate from time $t-1$ to time t is defined as

$$\Delta d_t = \log\left(\frac{D_t}{D_{t-1}}\right) = \delta_{0X} + \delta_{1X}' X_t,$$

where δ_{0X} is a scalar and δ_{1X} is an $N \times 1$ vector.

Given the above equations, we can derive the prices of zero-coupon bonds, zero-coupon equities and the aggregate market index as shown in Chapter 1. To estimate the asset prices, I use the result of observational equivalence between the latent state vector X_t and a set of observable state factors $\mathcal{P}_t$ as shown in JSZ. The same method was used in Chapter 1 to price equities. However, because the aim of this chapter is to price both stocks and bonds, a different set of observable state factors $\mathcal{P}_t$ is used here. Specifically, $\mathcal{P}_t$ consists of portfolios of yields PC_t and dividend growth, i.e.

$$\mathcal{P}_t = \begin{bmatrix} \mathrm{PC}_t \\ \Delta d_t \end{bmatrix}.$$

PC_t denotes the principal components of bond and equity yields (portfolio of yields) and is a $(N-1)\times 1$ vector. More specifically,

$$\mathrm{PC}_t = W y_t.$$

Here y_t is a $J \times 1$ vector and includes all the bond and equity yields at time t. W is $(N-1)\times J$ and is the weight of the portfolios. Therefore, given that y_t is affine in X_t,

$$y_t = A_X + B_X' X_t,$$

where A_X is $J \times 1$ and B_X is $N \times J$, $\mathcal{P}_t$ is also affine in X_t as

$$\mathcal{P}_t = \begin{bmatrix} \mathrm{PC}_t \\ \Delta d_t \end{bmatrix} = \begin{bmatrix} W A_X \\ \delta_{0X} \end{bmatrix} + \begin{bmatrix} W B_X' \\ \delta_{1X}' \end{bmatrix} X_t = A + B X_t.$$

Given the affine relationship between $\mathcal{P}_t$ and X_t, we can show that the economy can be observationally equivalently defined in terms of $\mathcal{P}_t$ as follows:

$$\begin{aligned} \Delta \mathcal{P}_t &= K_{0\mathcal{P}}^{\mathbb{P}} + K_{1\mathcal{P}}^{\mathbb{P}} \mathcal{P}_{t-1} + \Sigma_{\mathcal{P}} \epsilon_t^{\mathbb{P}}, \\ \Delta \mathcal{P}_t &= K_{0\mathcal{P}}^{\mathbb{Q}} + K_{1\mathcal{P}}^{\mathbb{Q}} \mathcal{P}_{t-1} + \Sigma_{\mathcal{P}} \epsilon_t^{\mathbb{Q}}, \\ r_t &= \rho_{0\mathcal{P}} + \rho_{1\mathcal{P}}' \mathcal{P}_t, \\ \Delta d_t &= \delta_{0\mathcal{P}} + \delta_{1\mathcal{P}}' \mathcal{P}_t. \end{aligned}$$

This result is shown in Proposition 2 in Chapter 1. And the estimation can be carried out in a two-step procedure. First, the $\mathbb{P}$ dynamics of the state process can be estimated by an unrestricted VAR. Then, taking the estimated parameters from the VAR as given, the rest of the parameters are then estimated using MLE.

2.3.1 Data

This section describes the data used in the estimation, including bond yields and the principal components extracted from both zero-coupon bond yields and zero-coupon equity yields. The data on equity yields and dividend growth used in this chapter is the same as in Chapter 1. Previously, there was no empirical counterpart for zero-coupon equities, so in the literature the estimation of the co-pricing model, e.g. in Bekaert and Grenadier (2001), can only use the aggregate equity price, which is the sum of the prices of zero-coupon equities of all maturities. Hence, the information on the term structure of equities was missing in the estimation. However, Chapter 1 shows that, using dividend strip prices, the S&P 500 index and the dividend series provided by van Binsbergen, Brandt and Koijen (2012), we can construct "equity yields" that are comparable to "bond yields".

Bond yields

For bonds, end-of-month Constant Maturity Treasury yields with maturities of 6 months, 1, 2, 3, 5, 7 and 10 years are taken from February 1996 to October 2009. I also include the one-month interest rate provided by BBK as an additional bond yield. The time series of the monthly zero-coupon bond yields are plotted in Figure 2.1. The mean and standard deviation of the yields are shown in Table 2.1. From the figure and the table, we can see that bond yields for the sample period exhibit some stylized facts. The mean bond yield curve is upward sloping. Standard deviations of bond yields generally decrease with maturity. Finally, yields are highly autocorrelated, with increasing autocorrelation at longer maturities. In dealing with the stationarity of bond yields at various maturities, I follow the same reasoning as in Chapter 1 for the bond yield of one-month maturity. That is, I assume all bond yields are stationary, but allow them to have a very low speed of mean reversion.

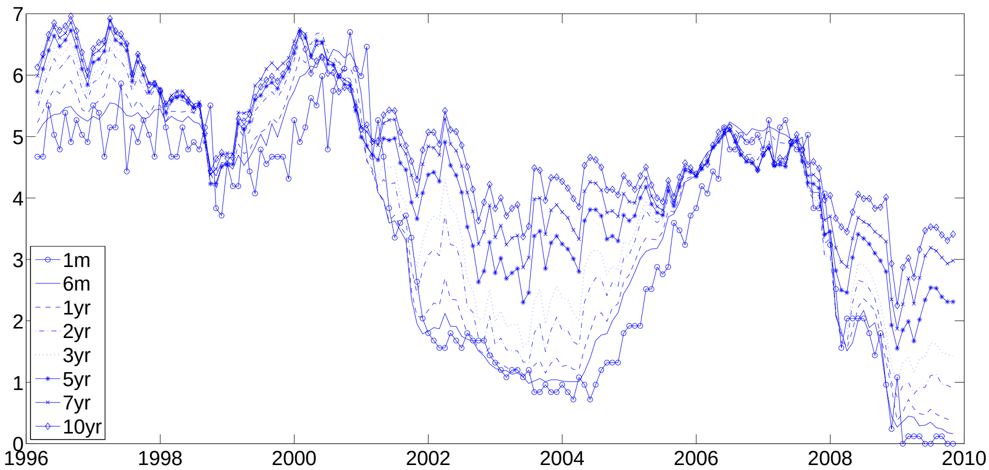


Figure 2.1: Zero-coupon bond yields, 1996m2:2009m10

Table 2.1: Summary statistics of the U.S. bond yields (all numbers are in annualized percentages)

Maturity (years)	1m	6m	1y	2y	3y	5y	7y	10y
Panel A: Data								
Mean	3.37	3.57	3.69	3.95	4.13	4.45	4.71	4.87
Standard deviation	1.87	1.90	1.84	1.74	1.60	1.33	1.19	1.02

Co-pricing factors

There are eight zero-coupon bond yields and five zero-coupon equity yields available monthly from February 1996 to October 2009. To estimate the model, five principal components are extracted from these thirteen series. Here, five PCs are chosen due to the fact that existing factor models on co-pricing such as Lettau and Wachter (2011), Koijen, Lustig and Van Nieuwerburgh (2013) and Ang and Ulrich (2012) all happen to use six state factors to drive their economy. If a six-factor state vector is also employed here, then the state vector in any of the papers mentioned above can be viewed as a rotation of the state vector in this chapter. Hence it is easier to translate between each paper's results. More importantly, whereas existing papers on co-pricing set many additional restrictions on top of the minimal set of restrictions imposed by the maximal identifiable model, we can impose the same set of restrictions on the estimated maximal identifiable model in this chapter, and see how these additional restrictions affect asset pricing moments. Hence, five PCs are chosen and, together with monthly dividend growth for the same sample period taken from van Binsbergen, Brandt and Koijen (2012), they make up the six state factors that drive the economy and all asset prices.

2.3.2 Estimation results

This subsection shows the estimation results and compares the data with the model-implied asset pricing moments of zero-coupon bonds, zero-coupon equities, dividend strips and the aggregate stock market index.

Model parameters

($\lambda^{\mathbb{Q}}$, $r^{\mathbb{Q}}_{\infty}$, $\Sigma_{\mathcal{P}}$, δ_{0X}, δ_{1X}) are estimated in MLE and their estimates are shown in Table 2.2 and Table 2.3. From the estimates of $(\lambda^{\mathbb{Q}}_i + 1)$s, we can see that the estimated state process has real and distinct eigenvalues, and because all $(\hat{\lambda}^{\mathbb{Q}}_i + 1)$s are between zero and one the estimated state process is stationary under $\mathbb{Q}$. Moreover, the estimate of the risk-neutral long-run mean of the short rate is now 6.68% per annum, reflecting the fact that by adding bond yield data we can achieve a more reasonable estimate of the short rate parameter than using only equity data. Table 2.4 reports the OLS estimates of $K_{0\mathcal{P}}^{\mathbb{P}}$ and $K_{1\mathcal{P}}^{\mathbb{P}}+I$. Estimates are in annual numbers and their standard deviations are provided in parentheses.

Table 2.2: Maximum likelihood estimates of the risk-neutral parameters, unrestricted co-pricing model

	$\lambda_1^{\mathbb{Q}}+1$	$\lambda_2^{\mathbb{Q}}+1$	$\lambda_3^{\mathbb{Q}}+1$	$\lambda_4^{\mathbb{Q}}+1$	$\lambda_5^{\mathbb{Q}}+1$	$\lambda_6^{\mathbb{Q}}+1$	$r_\infty^{\mathbb{Q}}$
Estimate	0.9967	0.9930	0.9358	0.2799	0.2779	0.2761	6.6794
Standard deviation	(0.0010)	(0.0007)	(0.0031)	(0.0019)	(3.6189e-5)	(0.0017)	(0.5841)
	δ_{0X}	$\delta_{1X,1}$	$\delta_{1X,2}$	$\delta_{1X,3}$	$\delta_{1X,4}$	$\delta_{1X,5}$	$\delta_{1X,6}$
Estimate	0.0185	−0.0019	0.7058	0.4636	−0.5838	0.0083	0.5380
Standard deviation	(0.0118)	(1.6274e-5)	(0.0012)	(0.0089)	(0.0013)	(0.0001)	(0.0047)

Table 2.3: Maximum likelihood estimates of the conditional covariance, unrestricted co-pricing model

	$\Sigma_{\mathcal{P},11}$					
Estimate	21.0251					
Standard deviation	(0.1593)					
	$\Sigma_{\mathcal{P},21}$	$\Sigma_{\mathcal{P},22}$				
Estimate	−0.2952	0.5408				
Standard deviation	(0.0015)	(0.0061)				
	$\Sigma_{\mathcal{P},31}$	$\Sigma_{\mathcal{P},32}$	$\Sigma_{\mathcal{P},33}$			
Estimate	0.0383	−0.2494	0.2995			
Standard deviation	(0.0001)	(0.0037)	(0.0043)			
	$\Sigma_{\mathcal{P},41}$	$\Sigma_{\mathcal{P},42}$	$\Sigma_{\mathcal{P},43}$	$\Sigma_{\mathcal{P},44}$		
Estimate	0.1310	0.0836	−0.0658	0.1194		
Standard deviation	(0.0012)	(0.0023)	(0.0001)	(0.0041)		
	$\Sigma_{\mathcal{P},51}$	$\Sigma_{\mathcal{P},52}$	$\Sigma_{\mathcal{P},53}$	$\Sigma_{\mathcal{P},54}$	$\Sigma_{\mathcal{P},55}$	
Estimate	−0.0507	−0.0234	0.1276	−0.0239	0.2783	
Standard deviation	(0.0002)	(0.0001)	(0.0010)	(0.0002)	(0.0025)	
	$\Sigma_{\mathcal{P},61}$	$\Sigma_{\mathcal{P},62}$	$\Sigma_{\mathcal{P},63}$	$\Sigma_{\mathcal{P},64}$	$\Sigma_{\mathcal{P},65}$	$\Sigma_{\mathcal{P},66}$
Estimate	20.6796	3.7272	1.0947	17.9939	11.3924	45.9478
Standard deviation	(0.6679)	(0.0294)	(0.0033)	(0.0384)	(0.0361)	(0.4947)

Table 2.4: Maximum likelihood estimates of the physical parameters, unrestricted co-pricing model

	$K^{\mathbb{P}}_{0\mathcal{P},1}$	$K^{\mathbb{P}}_{0\mathcal{P},2}$	$K^{\mathbb{P}}_{0\mathcal{P},3}$	$K^{\mathbb{P}}_{0\mathcal{P},4}$	$K^{\mathbb{P}}_{0\mathcal{P},5}$	$K^{\mathbb{P}}_{0\mathcal{P},6}$
Estimate	0.2898	0.8491	−0.2154	0.2318	0.4301	42.2813
Standard deviation	(15.7761)	(0.4531)	(0.2819)	(0.1353)	(0.2344)	(41.1135)
	$K^{\mathbb{P}}_{1\mathcal{P},11}+1$	$K^{\mathbb{P}}_{1\mathcal{P},12}$	$K^{\mathbb{P}}_{1\mathcal{P},13}$	$K^{\mathbb{P}}_{1\mathcal{P},14}$	$K^{\mathbb{P}}_{1\mathcal{P},15}$	$K^{\mathbb{P}}_{1\mathcal{P},16}$
Estimate	0.7730	0.6810	−0.7070	−3.8695	4.3913	−0.0429
Standard deviation	(0.0484)	(0.3778)	(1.8274)	(3.4290)	(4.6373)	(0.0264)
	$K^{\mathbb{P}}_{1\mathcal{P},21}$	$K^{\mathbb{P}}_{1\mathcal{P},22}+1$	$K^{\mathbb{P}}_{1\mathcal{P},23}$	$K^{\mathbb{P}}_{1\mathcal{P},24}$	$K^{\mathbb{P}}_{1\mathcal{P},25}$	$K^{\mathbb{P}}_{1\mathcal{P},26}$
Estimate	−0.0043	0.9948	−0.0326	−0.0621	−0.5000	0.0009
Standard deviation	(0.0014)	(0.0109)	(0.0525)	(0.0985)	(0.1332)	(0.0008)
	$K^{\mathbb{P}}_{1\mathcal{P},31}$	$K^{\mathbb{P}}_{1\mathcal{P},32}$	$K^{\mathbb{P}}_{1\mathcal{P},33}+1$	$K^{\mathbb{P}}_{1\mathcal{P},34}$	$K^{\mathbb{P}}_{1\mathcal{P},35}$	$K^{\mathbb{P}}_{1\mathcal{P},36}$
Estimate	0.0001	0.0082	0.8817	0.1298	−0.5027	−0.0003
Standard deviation	(0.0009)	(0.0068)	(0.0326)	(0.0613)	(0.0829)	(0.0005)
	$K^{\mathbb{P}}_{1\mathcal{P},41}$	$K^{\mathbb{P}}_{1\mathcal{P},42}$	$K^{\mathbb{P}}_{1\mathcal{P},43}$	$K^{\mathbb{P}}_{1\mathcal{P},44}+1$	$K^{\mathbb{P}}_{1\mathcal{P},45}$	$K^{\mathbb{P}}_{1\mathcal{P},46}$
Estimate	−0.0012	−0.0006	0.0154	0.9129	0.1006	−0.0000
Standard deviation	(0.0004)	(0.0032)	(0.0157)	(0.0294)	(0.0398)	(0.0002)
	$K^{\mathbb{P}}_{1\mathcal{P},51}$	$K^{\mathbb{P}}_{1\mathcal{P},52}$	$K^{\mathbb{P}}_{1\mathcal{P},53}$	$K^{\mathbb{P}}_{1\mathcal{P},54}$	$K^{\mathbb{P}}_{1\mathcal{P},55}+1$	$K^{\mathbb{P}}_{1\mathcal{P},56}$
Estimate	0.0010	0.0020	−0.0260	0.0668	0.4512	−0.0006
Standard deviation	(0.0007)	(0.0056)	(0.0271)	(0.0509)	(0.0689)	(0.0004)
	$K^{\mathbb{P}}_{1\mathcal{P},61}$	$K^{\mathbb{P}}_{1\mathcal{P},62}$	$K^{\mathbb{P}}_{1\mathcal{P},63}$	$K^{\mathbb{P}}_{1\mathcal{P},64}$	$K^{\mathbb{P}}_{1\mathcal{P},65}$	$K^{\mathbb{P}}_{1\mathcal{P},66}+1$
Estimate	−0.5201	1.1313	−1.7847	−16.8210	6.7022	−0.2151
Standard deviation	(0.1261)	(0.9846)	(4.7622)	(8.9361)	(12.0850)	(0.0687)

Equities

a. Equity yields

Table 2.5 shows that the model's predictions regarding yields of zero-coupon equities can match the data well in terms of mean and standard deviation, just like in Chapter 1.

Table 2.5: Estimation results: moments comparison for equity yields, annual percentages, unrestricted co-pricing model

Maturity	6m	12m	18m	24m	50y
Panel A: Data					
Mean	6.91	4.38	3.53	3.11	1.85
Standard deviation	28.60	14.17	9.36	6.96	0.38
Panel B: Model					
Mean	6.94	4.33	3.50	3.11	1.89
Standard deviation	28.65	14.13	9.30	6.90	0.38

b. Return of the dividend strip and the aggregate market

Just as in Chapter 1, we calculate the return of the dividend strip of short maturity and the return of the aggregate market to see whether the slope of the equity term structure is downward sloping or upward sloping. However, in this chapter, the slope of the equity term structure is backed out from the joint estimation of both the bond yield curve and the equity yield curve. Hence it is interesting to see that, after adding bond data, the model-implied slope of the equity term structure can still maintain its downward slope given an upward sloping bond yield curve.

Table 2.6 repeats for convenience Table 1.5 in Chapter 1, which shows the return of the short-term dividend strip of average maturity between 1.3 years and 1.9 years, the return of the S&P 500 index and their returns in excess of the short rate.

Table 2.6: Data: summary statistics of the return of the short-term strip and the market

	$R_{\text{short strip}}$	$R_{\text{short strip}} - R_f$	R_{market}	$R_{\text{market}} - R_f$
Mean	13.90	10.53	6.67	3.29
Standard deviation	27.03	27.04	16.26	16.22
Sharpe ratio	0.39	—	0.20	—

The returns of the dividend strip and of the market index are calculated using the equations below:

$$R^{\mathrm{s}}_{t+1} = \frac{P^{\mathrm{s}}_{n-1,t+1} + D_{t+1}}{P^{\mathrm{s}}_{n,t}} - 1,$$

where $P^{\mathrm{s}}_{n,t} = \sum_{i=1}^{n} P^{\mathrm{d}}_{i,t}$ is the time-t price of the dividend strip of maturity n, the sum of the prices of the zero-coupon equities from maturity 1 to n;

$$R^{\mathrm{m}}_{t+1} = \frac{P^{\mathrm{m}}_{t+1} + D_{t+1}}{P^{\mathrm{m}}_{t}} - 1,$$

where $P^{\mathrm{m}}_{t} = \sum_{i=1}^{\infty} P^{\mathrm{d}}_{i,t}$ is the time-t market index, the sum of prices of zero-coupon equities from maturity 1 to ∞.

Table 2.7 and Table 2.8 show that the model-implied equity term structure is downward sloping, just as observed in the data. The returns and excess returns of the short-term dividend strips of maturities of fifteen months to twenty-three months are slightly lower than the returns of the short-term dividend strip, but they are comparable. Moreover the market return and excess market return are both closely matched.

Table 2.7: Summary statistics of the return of the short-term strip and the market, unrestricted co-pricing model

	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	11.90	11.76	11.64	11.54	11.44	11.36	11.28	11.21	11.15	7.05
Standard deviation	26.48	26.57	26.65	26.72	26.78	26.84	26.89	26.93	26.97	24.72
Sharpe ratio	0.32	0.32	0.31	0.31	0.30	0.30	0.29	0.29	0.29	0.15

Table 2.8: Summary statistics of the excess return of the short-term strip and the market, unrestricted co-pricing model

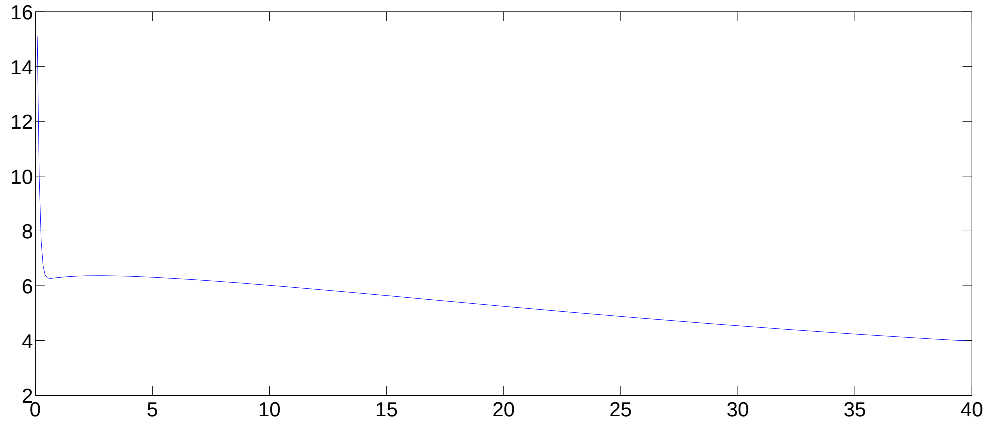
	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	8.55	8.42	8.30	8.19	8.10	8.01	7.93	7.86	7.80	3.70
Standard deviation	26.50	26.59	26.67	26.74	26.80	26.86	26.91	26.95	26.99	24.73

c. Risk premiums of zero-coupon equities

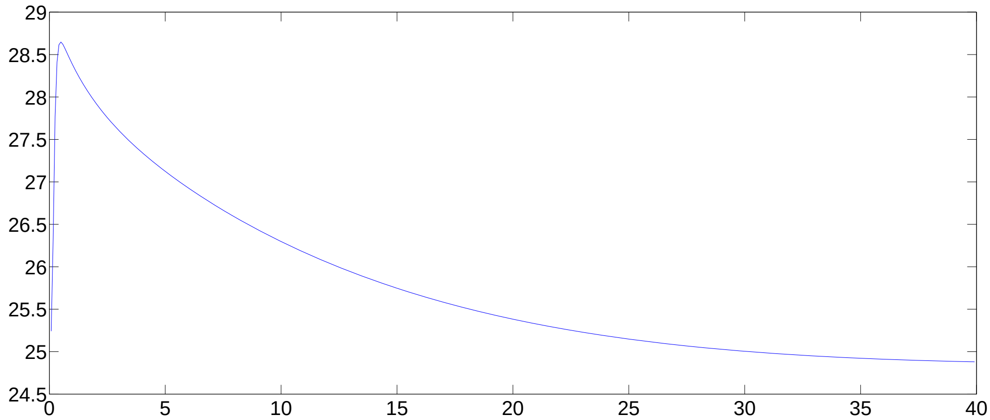
Using

$$R^{\mathrm{d}}_{n,t+1} - R_f = \frac{P^{\mathrm{d}}_{n-1,t+1}}{P^{\mathrm{d}}_{nt}} - R_f$$

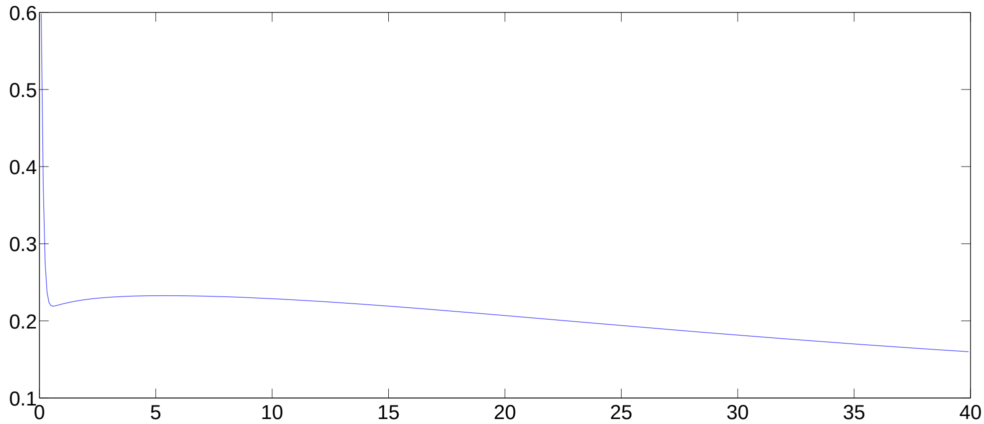
we calculate the risk premiums of the zero-coupon equities to see whether they follow the downward sloping pattern as in Lettau and Wachter (2007), and the answer is affirmative. The estimated risk premiums are strictly decreasing in maturity and the estimated volatilities and Sharpe ratios are generally decreasing in maturity, which are consistent with the results in Lettau and Wachter (2007). These are shown in Figure 2.2.



(a) Estimated annual mean risk premiums for ZCEs, maturity in years



(b) Estimated annual volatility of risk premiums for ZCEs, maturity in years



(c) Estimated annual Sharpe ratios of risk premium for ZCEs, maturity in years

Figure 2.2: Estimated risk premiums of zero-coupon equities, unrestricted co-pricing model

Bond yields

The estimation results regarding bonds are shown in Table 2.9. The table shows that model-implied bond yields exhibit the same characteristics as the observed yields, i.e. bond yields means are almost monotonically increasing with maturity, while the standard deviations of bond yields are generally decreasing in maturity. The close match in magnitudes of both the means and the standard deviations between the estimated and the observed bond yields shows that the co-pricing framework can match bond yields well. Hence this shows that the co-pricing framework in this paper can generate both the downward sloping equity yield curve and the upward sloping bond yield curve.

Table 2.9: Estimation results: moments comparison for zero-coupon bonds (all numbers are in annualized percentage), unrestricted co-pricing model

Maturity	1m	6m	1y	2y	3y	5y	7y	10y
Panel A: Data								
Mean	3.37	3.57	3.69	3.95	4.13	4.45	4.71	4.87
Standard deviation	1.87	1.90	1.84	1.74	1.60	1.33	1.19	1.02
Panel B: Model								
Mean	3.36	3.59	3.70	3.93	4.12	4.43	4.67	4.94
Standard deviation	1.84	1.98	1.86	1.67	1.54	1.35	1.21	1.04

In summary, using the maximal identifiable model and using data on both zero-coupon bonds and zero-coupon equities, the estimated model can generally match the time-series and cross-sectional properties of asset pricing moments for both bonds and equities.

2.4 Restricted Model

2.4.1 Comparison with Lettau and Wachter (2011)

This section illustrates the importance of taking into account the restrictions implied by the maximal identifiable model. I will use Lettau and Wachter (2011), one of the existing papers of co-pricing stocks and bonds, as an example to show that it is important to take into account the restrictions in order to match all the asset pricing moments.

The model in this paper and the model in Lettau and Wachter (2011) are both designed to price bonds and equities and, in particular, to simultaneously generate an upward sloping bond yield curve and a downward sloping equity yield curve. Both papers belong to the same class of models, i.e. Gaussian affine term structure models, and have the same number of factors. Hence, once they are normalized to the same canonical form, the resulting two models should have very similar state processes. However, in Lettau and Wachter (2011), the model is calibrated rather than estimated and they implicitly impose additional restrictions on their canonical form. For example, in their model, one of the assumptions is that only dividend risk is priced directly and hence the price of risk matrix reduces to a single time varying vector and the time variation in risk premiums depends only on a one-dimensional state variable. Hence it will be interesting to see how this restriction affects asset-pricing predictions. The estimation results of the restricted model are shown below.

2.4.2 Estimation results

This subsection shows the estimation results of the restricted model and compares them with the estimation results of the unrestricted model.

Model parameters

Table 2.10, Table 2.11 and Table 2.12 list the estimates of the model parameters. From the estimates of the $(\lambda_i^{\mathbb{Q}} + 1)$s, we can see that the restricted state process is also stationary under $\mathbb{Q}$.

Table 2.10: Maximum likelihood estimates of the risk-neutral parameters, restricted co-pricing model

	$\lambda_1^{\mathbb{Q}}+1$	$\lambda_2^{\mathbb{Q}}+1$	$\lambda_3^{\mathbb{Q}}+1$	$\lambda_4^{\mathbb{Q}}+1$	$\lambda_5^{\mathbb{Q}}+1$	$\lambda_6^{\mathbb{Q}}+1$	$r_\infty^{\mathbb{Q}}$
Estimate	0.9951	0.9919	0.9496	0.2793	0.2772	0.2753	6.2736
Standard deviation	(0.0015)	(0.0148)	(0.0324)	(0.0076)	(0.0001)	(0.0065)	(7.1347)
	δ_{0X}	$\delta_{1X,1}$	$\delta_{1X,2}$	$\delta_{1X,3}$	$\delta_{1X,4}$	$\delta_{1X,5}$	$\delta_{1X,6}$
Estimate	0.0192	−0.0020	0.7024	0.4603	−0.5866	0.0078	0.5379
Standard deviation	(0.0645)	(0.0007)	(0.0356)	(0.0108)	(0.0377)	(0.0014)	(0.0057)

Table 2.11: Maximum likelihood estimates of the conditional covariance, restricted co-pricing model

	$\Sigma_{\mathcal{P},11}$					
Estimate	22.1036					
Standard deviation	(0.4268)					
	$\Sigma_{\mathcal{P},21}$	$\Sigma_{\mathcal{P},22}$				
Estimate	−0.2913	0.5159				
Standard deviation	(0.0177)	(0.0203)				
	$\Sigma_{\mathcal{P},31}$	$\Sigma_{\mathcal{P},32}$	$\Sigma_{\mathcal{P},33}$			
Estimate	0.0242	−0.2276	0.2945			
Standard deviation	(0.0016)	(0.0191)	(0.0309)			
	$\Sigma_{\mathcal{P},41}$	$\Sigma_{\mathcal{P},42}$	$\Sigma_{\mathcal{P},43}$	$\Sigma_{\mathcal{P},44}$		
Estimate	0.1003	0.0585	−0.0636	0.1225		
Standard deviation	(0.0019)	(0.0012)	(0.0002)	(0.0099)		
	$\Sigma_{\mathcal{P},51}$	$\Sigma_{\mathcal{P},52}$	$\Sigma_{\mathcal{P},53}$	$\Sigma_{\mathcal{P},54}$	$\Sigma_{\mathcal{P},55}$	
Estimate	−0.0519	−0.0304	0.1291	−0.0237	0.2756	
Standard deviation	(0.0007)	(0.0042)	(0.0006)	(0.0017)	(0.0403)	
	$\Sigma_{\mathcal{P},61}$	$\Sigma_{\mathcal{P},62}$	$\Sigma_{\mathcal{P},63}$	$\Sigma_{\mathcal{P},64}$	$\Sigma_{\mathcal{P},65}$	$\Sigma_{\mathcal{P},66}$
Estimate	20.3495	3.7084	1.0906	17.8574	11.3552	48.2150
Standard deviation	(0.1923)	(0.0113)	(0.0946)	(0.0472)	(1.1370)	(2.3303)

Table 2.12: Maximum likelihood estimates of the physical parameters, restricted co-pricing model

	$K^{\mathbb{P}}_{0\mathcal{P},1}$	$K^{\mathbb{P}}_{0\mathcal{P},2}$	$K^{\mathbb{P}}_{0\mathcal{P},3}$	$K^{\mathbb{P}}_{0\mathcal{P},4}$	$K^{\mathbb{P}}_{0\mathcal{P},5}$	$K^{\mathbb{P}}_{0\mathcal{P},6}$
Estimate	7.4364	−0.0688	0.1718	0.0090	0.3132	47.2161
Standard deviation	(20.7115)	(8.7557)	(2.6950)	(2.7586)	(0.8602)	(82.3956)
	$K^{\mathbb{P}}_{1\mathcal{P},11}+1$	$K^{\mathbb{P}}_{1\mathcal{P},12}$	$K^{\mathbb{P}}_{1\mathcal{P},13}$	$K^{\mathbb{P}}_{1\mathcal{P},14}$	$K^{\mathbb{P}}_{1\mathcal{P},15}$	$K^{\mathbb{P}}_{1\mathcal{P},16}$
Estimate	0.7791	0.4212	0.7357	−0.9317	−2.8522	−0.0303
Standard deviation	(0.0635)	(0.4960)	(2.3990)	(4.5017)	(6.0880)	(0.0346)
	$K^{\mathbb{P}}_{1\mathcal{P},21}$	$K^{\mathbb{P}}_{1\mathcal{P},22}+1$	$K^{\mathbb{P}}_{1\mathcal{P},23}$	$K^{\mathbb{P}}_{1\mathcal{P},24}$	$K^{\mathbb{P}}_{1\mathcal{P},25}$	$K^{\mathbb{P}}_{1\mathcal{P},26}$
Estimate	−0.0047	1.0065	−0.1062	0.1205	−0.5613	−0.0001
Standard deviation	(0.0269)	(0.2097)	(1.0142)	(1.9031)	(2.5737)	(0.0146)
	$K^{\mathbb{P}}_{1\mathcal{P},31}$	$K^{\mathbb{P}}_{1\mathcal{P},32}$	$K^{\mathbb{P}}_{1\mathcal{P},33}+1$	$K^{\mathbb{P}}_{1\mathcal{P},34}$	$K^{\mathbb{P}}_{1\mathcal{P},35}$	$K^{\mathbb{P}}_{1\mathcal{P},36}$
Estimate	−0.0007	0.0253	0.9684	0.0948	−0.6140	−0.0008
Standard deviation	(0.0083)	(0.0645)	(0.3122)	(0.5858)	(0.7922)	(0.0045)
	$K^{\mathbb{P}}_{1\mathcal{P},41}$	$K^{\mathbb{P}}_{1\mathcal{P},42}$	$K^{\mathbb{P}}_{1\mathcal{P},43}$	$K^{\mathbb{P}}_{1\mathcal{P},44}+1$	$K^{\mathbb{P}}_{1\mathcal{P},45}$	$K^{\mathbb{P}}_{1\mathcal{P},46}$
Estimate	−0.0013	−0.0008	−0.0097	0.9572	0.0932	−0.0004
Standard deviation	(0.0085)	(0.0661)	(0.3195)	(0.5996)	(0.8109)	(0.0046)
	$K^{\mathbb{P}}_{1\mathcal{P},51}$	$K^{\mathbb{P}}_{1\mathcal{P},52}$	$K^{\mathbb{P}}_{1\mathcal{P},53}$	$K^{\mathbb{P}}_{1\mathcal{P},54}$	$K^{\mathbb{P}}_{1\mathcal{P},55}+1$	$K^{\mathbb{P}}_{1\mathcal{P},56}$
Estimate	−0.0001	0.0298	0.0056	0.1161	0.2790	−0.0013
Standard deviation	(0.0026)	(0.0206)	(0.0996)	(0.1870)	(0.2529)	(0.0014)
	$K^{\mathbb{P}}_{1\mathcal{P},61}$	$K^{\mathbb{P}}_{1\mathcal{P},62}$	$K^{\mathbb{P}}_{1\mathcal{P},63}$	$K^{\mathbb{P}}_{1\mathcal{P},64}$	$K^{\mathbb{P}}_{1\mathcal{P},65}$	$K^{\mathbb{P}}_{1\mathcal{P},66}+1$
Estimate	−0.5925	3.2701	4.5274	−9.0586	−17.7648	−0.2270
Standard deviation	(0.2528)	(1.9732)	(9.5440)	(17.9088)	(24.2197)	(0.1376)

Equities

a. Equity yields

Table 2.13 shows that the restricted model's predictions regarding yields of zero-coupon equities can also match the data well in terms of mean and standard deviation. The abilities of the restricted model and the unrestricted model in matching the data are comparable.

Table 2.13: Estimation results: moments comparison for equity yields, annual percentages, restricted co-pricing model

Maturity	6m	12m	18m	24m	50y
Panel A: Data					
Mean	6.91	4.38	3.53	3.11	1.85
Standard deviation	28.60	14.17	9.36	6.96	0.38
Panel B: Model					
Mean	6.92	4.35	3.52	3.12	1.86
Standard deviation	28.65	14.14	9.30	6.89	0.31

b. Return of the dividend strip and the aggregate market

Table 2.14 and Table 2.15 show that the restricted model-implied equity term structure is also downward sloping just as observed in the data. Moreover, comparing with the unrestricted model, as the number of priced risk factors is restricted to one, the risk premium would load on the most volatile factor, which in the present model is dividend growth. As a result we would expect the return of the short-term asset to be higher in the restricted model. It turns out that the return of the short-term dividend strip has only improved marginally. However, in the restricted model, the return of the market is now even higher than in the data. Hence, although the restriction makes the return of the short-term dividend strip get closer to the high return of the short-term asset in the data, it has done so at the expense of a less downward sloping equity term structure.

Table 2.14: Summary statistics of the return of the short-term strip and the market, restricted co-pricing model

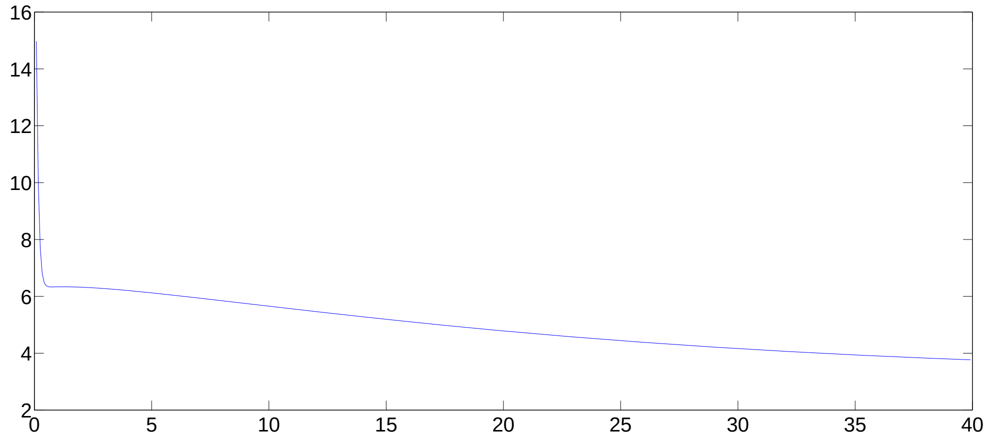
	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	11.93	11.79	11.67	11.56	11.47	11.38	11.30	11.23	11.16	7.41
Standard deviation	26.50	26.59	26.67	26.74	26.81	26.86	26.91	26.95	26.99	23.27
Sharpe ratio	0.32	0.32	0.31	0.31	0.30	0.30	0.30	0.29	0.29	0.17

Table 2.15: Summary statistics of the excess return of the short-term strip and the market, restricted co-pricing model

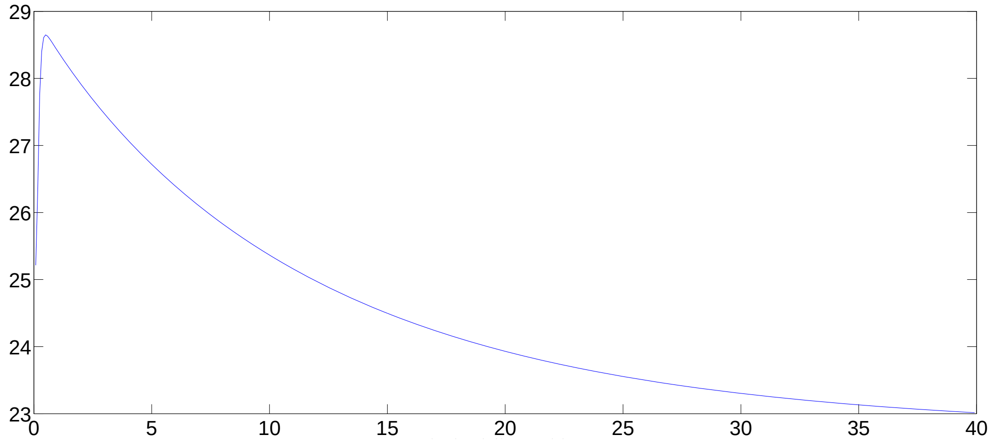
	R_{15m}	R_{16m}	R_{17m}	R_{18m}	R_{19m}	R_{20m}	R_{21m}	R_{22m}	R_{23m}	R_{market}
Mean	8.58	8.44	8.32	8.21	8.12	8.03	7.95	7.88	7.81	4.06
Standard deviation	26.52	26.61	26.69	26.76	26.82	26.88	26.93	26.97	27.01	23.29

c. Risk premiums of zero-coupon equities

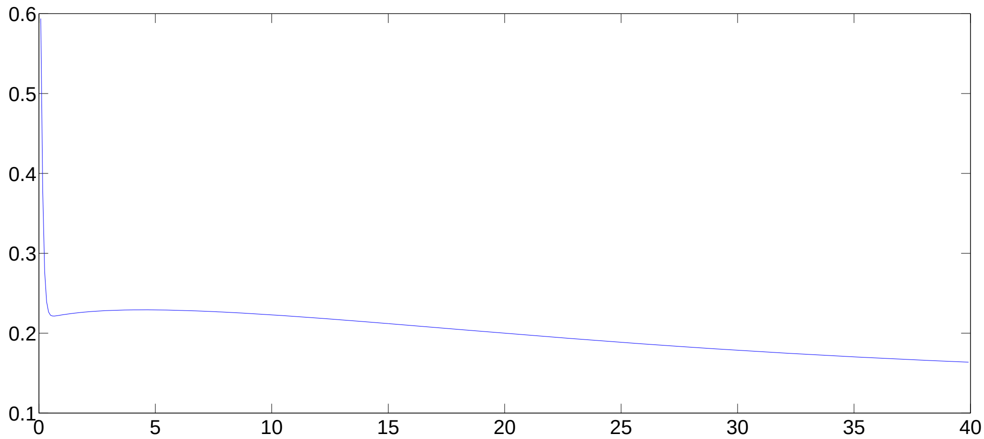
As before, the risk premiums of the zero-coupon equities of maturity up to forty years are plotted. Figure 2.3 shows that the estimated mean, volatility and Sharpe ratio of the risk premiums in the restricted model are very close to those in the unrestricted model. Hence, imposing additional restrictions does not affect the risk premiums of zero-coupon equities much.



(a) Estimated annual mean risk premiums for ZCEs, maturity in years



(b) Estimated annual volatility of risk premiums for ZCEs, maturity in years



(c) Estimated annual Sharpe ratios of risk premiums for ZCEs, maturity in years

Figure 2.3: Estimated risk premiums of zero-coupon equities, restricted co-pricing model

Bond yields

Table 2.16 shows the summary statistics of the bond yields implied by the restricted model. Bond yields implied by the restricted model exhibit the increasing mean and decreasing volatility in bond yield maturity that we observed in the data.

Table 2.16: Estimation results: moments comparison for zero-coupon bonds (all numbers are in annualized percentage), restricted co-pricing model

Maturity	1m	6m	1y	2y	3y	5y	7y	10y
Panel A: Data								
Mean	3.37	3.57	3.69	3.95	4.13	4.45	4.71	4.87
Standard deviation	1.87	1.90	1.84	1.74	1.60	1.33	1.19	1.02
Panel B: Model								
Mean	3.36	3.58	3.71	3.93	4.12	4.43	4.67	4.93
Standard deviation	1.85	1.95	1.86	1.68	1.55	1.35	1.21	1.03

In summary, we can see that imposing the same restriction, i.e. restricting the number of priced risk factors to one, makes the return of the short-term dividend strip get marginally closer to the high return of the short-term asset in the data. However, it will lead to a less downward sloping equity term structure. The overall performance of the model in matching the asset pricing moments is reduced. Hence this is an over-restriction that should not be imposed. Therefore, while the co-pricing models in general try to match many moments, it is important to impose a minimal number of restrictions on the model, i.e. to use a maximally identifiable model, to generally match all the asset pricing moments well without improving the fitness of some moments at the expense of others.

Future Research

In the current setup, stocks and bonds share the same shocks to the economy, which could potentially lead to tight co-movement between stocks and bonds. However, market frictions, investors' investment habits and regulatory rules could result in investors who only trade stocks and investors who only trade bonds. All of these are conditions could result in market segregation, reducing the co-movement between stocks and bonds. Hence the present model, where stocks and bonds share the same set of shocks, will overestimate the co-movement between stocks and bonds. A better model should incorporate the above-mentioned frictions.

2.5 Conclusion

This chapter estimates a maximal identifiable affine term structure model that explains the joint prices of stocks and bonds. Using the test assets of U.S. Treasury bonds and dividend strips, I show that the estimated model can generally match the time series and cross-sectional properties of zero-coupon bonds, zero-coupon equities and the aggregate stock index. Moreover, imposing restrictions prevalent in the co-pricing literature on the maximal model enhances certain features of the model, such as the high return of the short-term dividend strip, but reduces the model's ability to fit other aspects of the data, such as the level of the market risk premium.

Chapter 3

The Role of Asian Countries' Reserve Holdings on the International Yield Curves

3.1 Abstract

This chapter studies the effect of Asian countries' reserve holdings on the yield curves of six industrialized countries: the United States, the United Kingdom, Germany, Canada, Switzerland and Australia. A Gaussian affine term structure model with three yield factors and three unspanned macro factors including reserves is estimated to fit the yield curve of each country. Impulse responses and variance decompositions show that Asian countries' reserve holdings are an important factor affecting the international yield curves.

3.2 Introduction

In recent years, Asian countries have been accumulating their reserve holdings very rapidly. For example, the total foreign reserves of the three major holders of foreign reserves in Asia, namely China, Japan and South Korea, increased from \$113 billion at the beginning of 1990 to \$3.12 trillion at the beginning of 2009. A significant fraction of these reserves is believed to have been invested in the government bonds issued by several major industrialized countries, such as U.S. Treasury securities, U.K. Gilts and German Bunds. For example, at the end of July 2015, China held \$1.24 trillion in U.S. Treasury securities, and Japan held \$1.20 trillion in U.S. Treasury securities (as published on the U.S. Treasury website). Hence it is important to investigate how Asian countries' reserve holdings affect the yield curves across these countries.

The answer to the above question can have important implications for those countries' monetary policies, because if Asian reserves can significantly affect a country's yield curve, then the country's central banks must take Asian reserves into account in conducting its monetary policy. Failing to do so could result in monetary policy not reaching its intended goals. For example, in mid-2004, the Fed began to tighten monetary policy in order to raise interest rates. However, Greenspan (2005) noted that the longer-term interest rates failed to rise in response. Such a decoupling of long-term interest rates from the short-term interest rate is believed to have been caused by the strong purchase of U.S. Treasuries by Asian countries during this period. Therefore, we need to understand how a country's yield curve changes in response to Asian reserve holdings.

There exist a number of papers studying the effect that foreign countries' reserves or Treasuries holdings have on the U.S. yield curve. They mainly use two frameworks. The first framework is to regress bond yields on foreign countries' reserves holdings and other explanatory variables. For example, Beltran et al. (2013) first compute the term premium using an affine model without using reserve holdings, then regress the term premium on reserve holdings. Warnock and Warnock (2009) regress U.S. 10-year bond yields on standard macroeconomic variables as well as foreign official purchases of U.S. Treasury bonds. Sierra (2014) run a series of forecasting regressions of realized excess returns on measures of net purchases of Treasuries. However, this framework does not take into account the no-arbitrage condition when determining bond yields. Hence the effect of reserve holdings on certain bond yields included in the study may not apply to the whole bond yield curve.

The second framework uses affine term structure models to study the effect of reserves on bond yields, which imposes no-arbitrage condition to ensure bond yields are priced consistently. For example, Bernanke, Reinhart and Sack (2004) include observable economic and monetary variables into a no-arbitrage term structure model, and find that Treasury yields declined significantly during intervals around Japanese interventions to purchase dollars in the 2000-04 period. Rudebusch, Swanson and Wu (2006) estimate term structure models of Treasury yields including latent factors as state variables and find that foreign official holdings have no explanatory power. Apart from no-arbitrage, another important feature common to the second framework is the implicit assumption that macro variables included in the model are spanned by bond yields. This implies that, in this type of setup, reserves are often modeled as a risk factor that is spanned by bond yields, i.e. reserves can be expressed as a linear combination of bond yields. However, as pointed out by Joslin, Priebsch and Singleton (2014), when applied to a number of macro variables commonly included in macro-finance term structure models such macro-spanning assumptions are often strong and counterfactual. Moreover, including macro factors in addition to a number of latent or yield factors (yields' principal components) to price bonds contradicts the fact that almost all of the cross-sectional variation of bond yields can be explained by a small number of latent or yield factors.

This paper models Asian countries' reserve holdings together with other macro variables as unspanned by bond yields. Macro factors together with yield factors follow a VAR under the physical measure $\mathbb{P}$. However, only yield factors determine the pricing of bonds under the risk-neutral measure $\mathbb{Q}$. The identification of the risk-neutral parameters is ensured by the canonical form of the model proposed by Joslin, Singleton and Zhu (2011), which is maximally identifiable. Although macro factors do not enter into the pricing of contemporaneous bond yields, because macro factors and yield factors follow an unrestricted VAR under $\mathbb{P}$, macro factors can help to predict future yields. Hence impulse responses and variance decompositions implied from the VAR can be obtained. Following Ang and Piazzesi (2003), these are used to study the yield curve's response toward a shock to reserve holdings and to attribute the forecast variance of a particular yield to shocks to reserves. Moreover, not only the effect of Asian countries' reserve holdings on the U.S. yield curve is studied, because Asian countries also invest their reserves in the bond markets of other industrialized countries; reserves' effects on the yield curves of the U.K., Germany, Canada, Switzerland and Australia are also investigated. Asian countries' holdings of government debt issuance by these countries are not as clearly estimated as their holdings of U.S. Treasury securities. However, these are the countries whose currencies are identified in the International Monetary Fund's Currency Composition of Official Foreign Exchange Reserves (COFER) database alongside the U.S. dollar.[1] Hence, these countries are more likely to have their government debt held by Asian countries and have their government yield curves affected by the fluctuations of Asian countries' foreign exchange reserve holdings.

The same bond pricing model with unspanned macro risks is fitted to each country's yield curve. Impulse responses implied by the estimation results show that an initial one standard deviation shock to Asian reserve holdings can increase or decrease bond yields of the countries studied by up to 18 basis points (bp) during the first five years. Variance decompo-

[1] Germany is used as a proxy for the euro area.

sitions show that shocks to reserves contribute a significant proportion of the forecast variances of yields. For example, at the five-year-ahead forecast horizon, the highest proportions of variance explained by reserves can range from 2.12% to 6.78% for the countries studied. Moreover, the explanatory power of reserves generally increases with the forecast horizon across different countries and maturities. Such evidence suggests that Asian countries' reserve holdings are an important factor affecting the yield curves of the above industrialized countries.

The rest of the chapter is organized as follows. Section 3.3 describes the data used in this chapter. Section 3.4 specifies a Gaussian affine term structure model with yield factors and unspanned macro factors. Section 3.5 summarizes the estimation strategy and presents the estimation results for each country included in the study. Section 3.6 concludes.

3.3 Data

3.3.1 Bond yields

I use the same international bond yields dataset as in Wright (2011). Quarterly data on zero-coupon bond yields of fourteen maturities, ranging from one quarter to forty quarters, from the first quarter of 1990 to the first quarter of 2009, are taken to be the sample for the estimation. The sample's starting period is chosen because Asian countries' accumulation of reserves is a recent phenomenon (beginning around 1990). The sample period ends in 2009 to exclude the ramifications of the global financial crisis, such as central banks' quantitative easing programs. Bond yields are constructed using local currencies. A more thorough analysis could be carried out by rebasing yields to the same currency or by modeling exchange rate explicitly, for example, to jointly estimate the term structure of bond yields and the term structure of forward exchange rates, which is beyond the scope of this chapter. Although to some extent exchange rates' effect on yields is limited by the fact that China's currency was pegged to the U.S. dollar until 2005 and China has since been operating a managed floating exchange rate regime, and another motivation for China and Japan to accumulate large amounts of foreign reserves is to be able to intervene in the foreign exchange market as to keep their currencies at a stable value.

Figures 3.1–3.6 plot the yields of maturities of one, four and twenty quarters for the United States, the United Kingdom, Germany, Canada, Switzerland and Australia. The three maturities are chosen to represent the short end, middle and long end of a country's yield curve. Impulse responses and variance decompositions will also be based on the three yields.

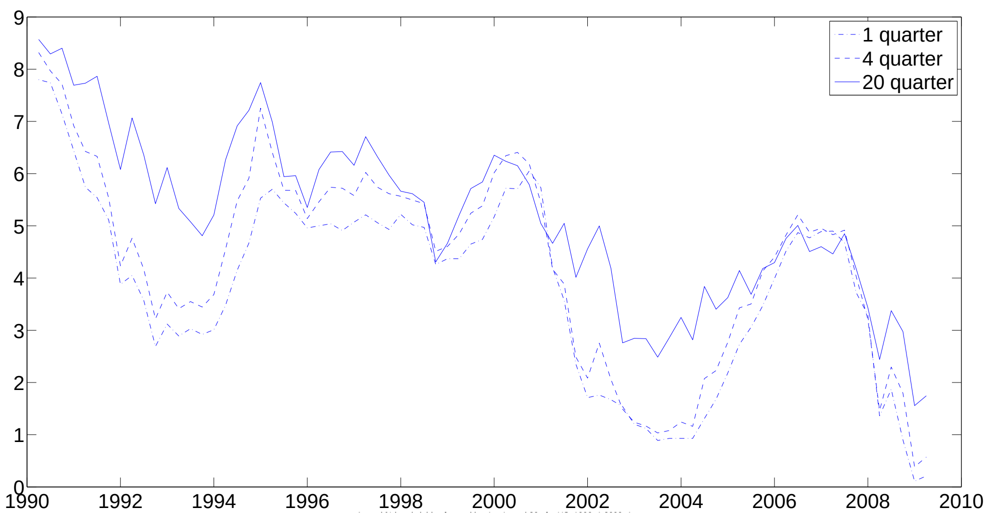


Figure 3.1: Annual % bond yields, United States

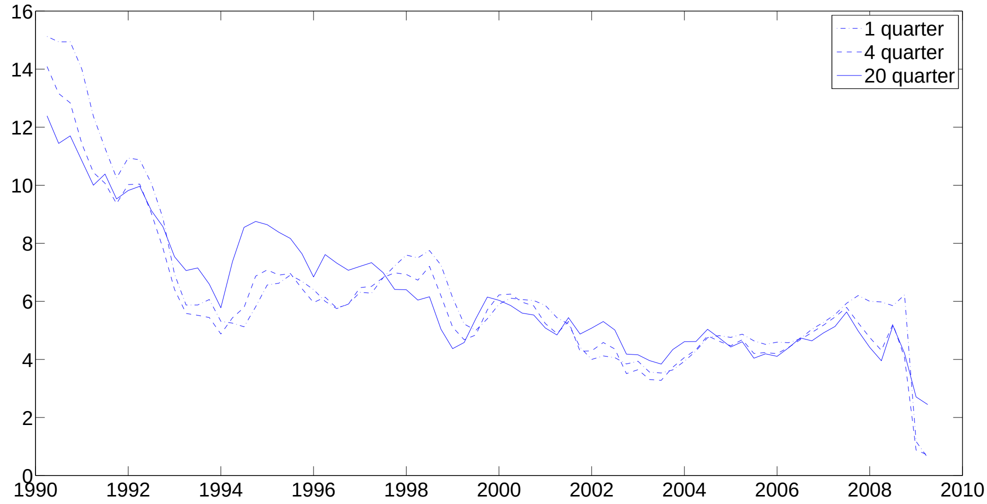


Figure 3.2: Annual % bond yields, United Kingdom

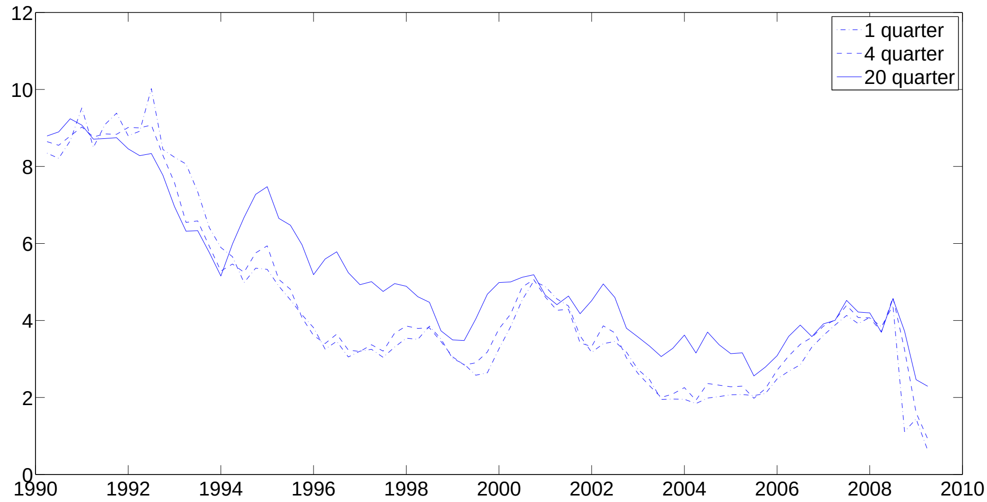


Figure 3.3: Annual % bond yields, Germany

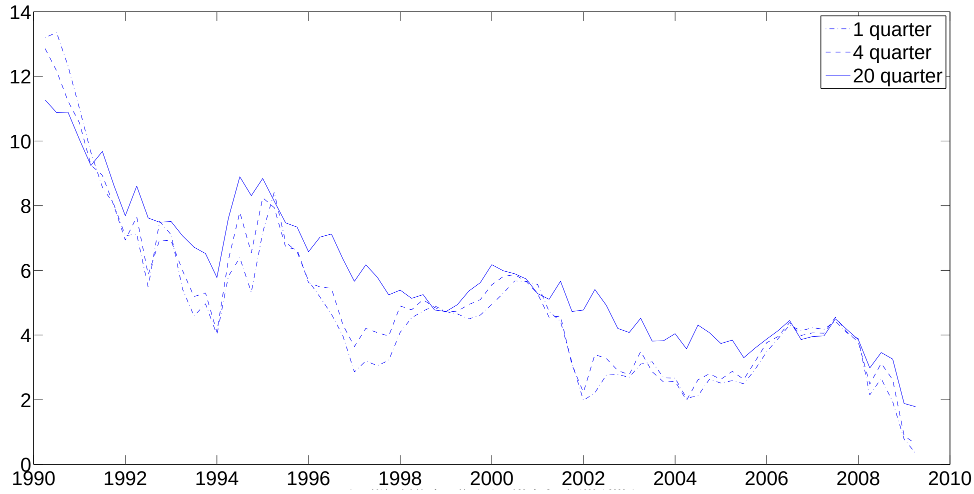


Figure 3.4: Annual % bond yields, Canada

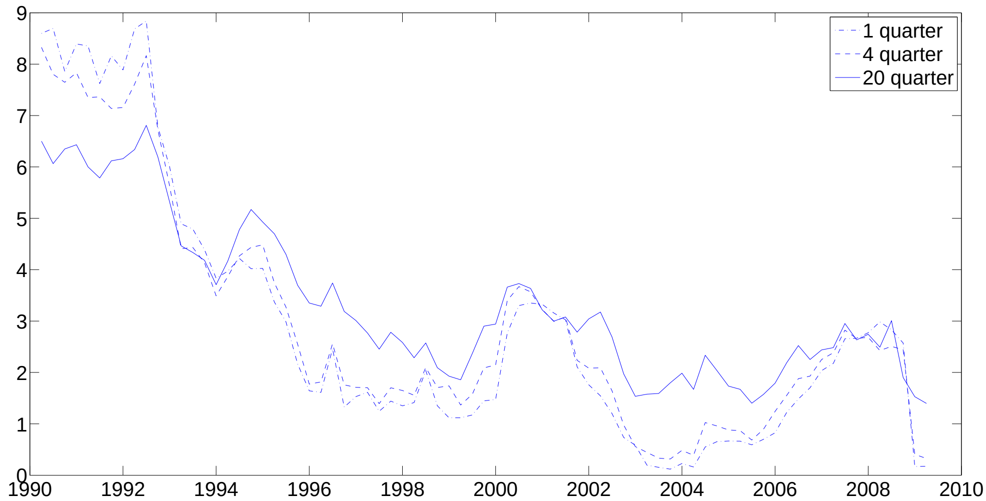


Figure 3.5: Annual % bond yields, Switzerland

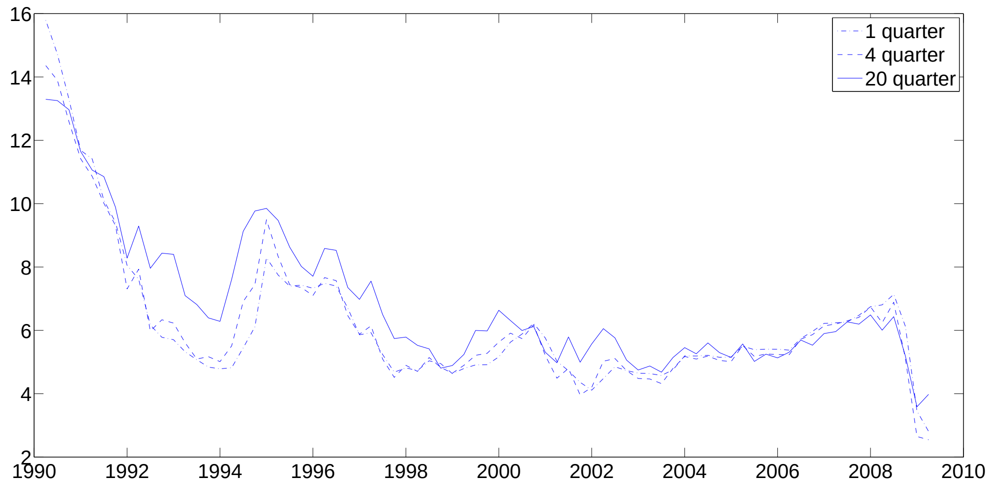


Figure 3.6: Annual % bond yields, Australia

Table 3.1 presents the summary statistics of the three yields presented above, along with some other maturities to show a more complete yield curve for each country. The table shows that, for the selected sample period, yields of most countries follow the pattern that the average bond yield is increasing in mean and bond volatility is decreasing in maturity. The yield curve of Germany closely tracks that of the United States. Yields of Canadian bonds are on average more than half a percentage higher than the yields of U.S. Treasuries. Swiss yields are low across all maturities, often being half of those of the other countries. Australia and the United Kingdom have higher yields than the other countries, but their yield curves are also flatter. For Australia, the difference between its one-quarter yield and twenty-quarter yield is only 60bp compared with 132bp for the United States. The one-quarter yield of the United Kingdom is higher than the rest of the maturities. For example, the one-quarter yield exceeds the four-quarter yield by 38bp, and exceeds the twenty-quarter yield by 7bp.

Table 3.1: Summary statistics of the international bond yields (all numbers are in annualized percentages)

Maturity (quarter)	1	4	8	12	20	28	40
United States							
Mean	3.85	4.32	4.59	4.81	5.17	5.46	5.79
Standard deviation	1.80	1.88	1.82	1.74	1.61	1.50	1.39
United Kingdom							
Mean	6.35	5.97	6.07	6.17	6.28	6.32	6.32
Standard deviation	2.79	2.40	2.27	2.21	2.18	2.17	2.14
Germany							
Mean	4.36	4.44	4.61	4.79	5.11	5.35	5.60
Standard deviation	2.33	2.16	2.06	1.98	1.84	1.73	1.62
Canada							
Mean	4.80	5.01	5.27	5.47	5.77	5.98	6.20
Standard deviation	2.57	2.39	2.25	2.18	2.10	2.04	2.03
Switzerland							
Mean	2.88	2.94	2.97	3.07	3.32	3.54	3.78
Standard deviation	2.52	2.23	1.90	1.71	1.52	1.41	1.31
Australia							
Mean	6.22	6.20	6.39	6.57	6.82	6.97	7.09
Standard deviation	2.30	2.18	2.14	2.13	2.14	2.16	2.18

3.3.2 Asian reserves

Since around 1990, Asian countries (especially those in East Asia) have started to accumulate reserves on an increasingly large scale. Figure 3.7 plots the time series of reserve holdings by China, Japan and Korea, the three largest holders of global reserves in the region. We can see that the accumulation of the three countries' reserves took off in 1994. At the end of 1993, China, Japan and Korea's reserve holdings were U.S.$21.8 billion, U.S.$98.4 billion and U.S.$20.5 billion, respectively. But by the beginning of 2009, the reserve holdings for the three countries reached U.S.$1929.5 billion, U.S.$990.1 billion and U.S.$203.1 billion, a total of more than U.S.$3 trillion. As a result of this accumulation, these countries' reserve holdings as a share of the global total reserves have increased significantly. For example, in 1990, the reserves of China were about 5% of world's total reserves. However, in 2009, China's reserves accounted for around 30% of the world's total reserves. Such an accumulation of reserves is mainly driven by the precautionary motives of these Asian countries. This is because, after the financial crises of the late 1990s, many emerging markets used reserves as a self-insurance against the volatility in the global financial market. And these reserves are often invested in the government bonds issued by the select few of the world's industrialized countries mentioned previously. Given the sheer magnitude of Asian reserves, it is necessary to understand their impact on the yield curves of those industrialized countries. And in this chapter, Asian countries' reserve holdings are modeled as a macro factor affecting the advanced economies' yield curves.

Moreover, as mentioned in the introduction to this chapter, Asian countries' reserve holdings will be modeled as a factor unspanned by the cross-section of bond yields, meaning that they do not enter into the risk-neutral pricing of bond yields, but they do affect bond risk premiums. Here, I provide empirical evidence for both. Table 3.2 lists the R^2 of regressing Asian reserves on a constant and on a country's yields based factors (the first three principal components, namely level, slope and curvature) for each country. We can see that the R^2 for all countries ranges from 30% to 40%, showing that Asian reserves are not linearly spanned by the bond yield curve of any country. Table 3.3 shows the results of regressing the excess returns of the ten-year bond on yield factors alone, and on both yield factors and macro factors including reserves. We can see that, for each country, including macros increases the R^2 over a model that only uses yield factors. Moreover, for each country's regression, "reserves" are significant at conventional levels, demonstrating the "reserves" value's ability to predict bond risk premiums. Therefore, both regression results direct us to model Asian reserves as an unspanned factor.

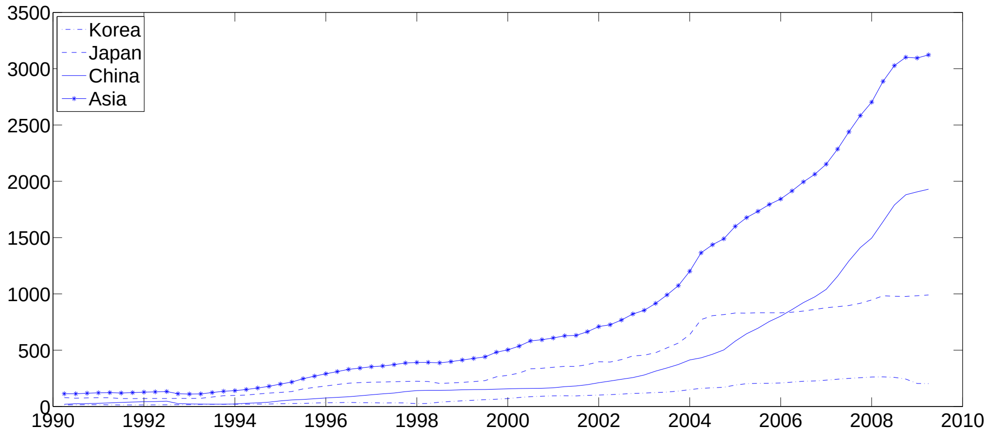


Figure 3.7: Reserve holdings of Asian countries, including China, Japan and South Korea, in billions of U.S. dollars

Table 3.2: R^2 of regressing reserves on yield curve factors for each country

	U.S.	U.K.	Germany	Canada	Switzerland	Australia
R^2	38.38%	30.80%	35.21%	39.25%	31.22%	40.98%

Table 3.3: Regressions of excess returns on ten-year bonds on yield factors and macro factors; Newey–West standard errors with four lags are shown in parentheses.

	U.S.		U.K.		Germany	
	Yields only	Include macros	Yields only	Include macros	Yields only	Include macros
Constant	0.0501	0.1117	0.0310	0.1813	0.0430	0.0341
	(0.0141)	(0.0743)	(0.0132)	(0.0768)	(0.0110)	(0.0569)
Level	0.3308	1.0663	0.4575	1.1325	0.1712	0.3719
	(0.1586)	(0.2529)	(0.0950)	(0.3052)	(0.1461)	(0.4689)
Slope	2.1116	2.9942	−0.8176	−1.0209	2.6396	2.9246
	(0.8021)	(0.6514)	(0.6627)	(0.5517)	(0.5824)	(0.7190)
Curvature	−1.8747	−1.1130	−2.0630	−1.8286	−3.4300	−3.6296
	(3.0548)	(3.2379)	(2.2669)	(2.3232)	(2.2237)	(1.9204)
GDP growth		−0.0112		−0.0295		−0.0064
		(0.0145)		(0.0168)		(0.0117)
Inflation		−0.0256		−0.0409		0.0022
		(0.0148)		(0.0179)		(0.0215)
Reserves		0.0010		0.0006		0.0004
		(0.0003)		(0.0003)		(0.0002)
R^2	22.79%	38.12%	26.64%	37.57%	35.83%	40.26%
	Canada		Switzerland		Australia	
	Yields only	Include macros	Yields only	Include macros	Yields only	Include macros
Constant	0.0528	0.0009	0.0320	0.0562	0.0464	0.0243
	(0.0127)	(0.0451)	(0.0102)	(0.0432)	(0.0170)	(0.0562)
Level	0.3044	0.4940	0.0829	0.8143	0.8406	0.9729
	(0.0902)	(0.0923)	(0.1661)	(0.3935)	(0.1050)	(0.1734)
Slope	−1.8442	−2.5611	3.7333	4.0073	−1.4647	−2.1210
	(0.8208)	(0.6057)	(0.7601)	(1.0459)	(1.2277)	(1.2312)
Curvature	−1.8410	−3.7623	−1.8078	−3.6883	−0.8432	0.3070
	(1.5983)	(2.0999)	(2.0759)	(1.8629)	(1.7427)	(2.2277)
GDP growth		−0.0009		0.0051		0.0018
		(0.0056)		(0.0114)		(0.0095)
Inflation		0.0105		−0.0286		−0.0014
		(0.0096)		(0.0146)		(0.0132)
Reserves		0.0008		0.0004		0.0005
		(0.0002)		(0.0002)		(0.0002)
R^2	25.14%	37.91%	31.64%	46.25%	38.25%	40.76%

3.3.3 Macro variables

The other two macro variables included in this chapter are GDP growth and inflation. Data on CPI inflation and GDP growth are from OECD's

Main Economic Indicators at the quarterly frequency. Both series have been smoothed with moving average filters to remove the seasonality. Figures 3.8–3.13 plot the time series of each country's annual percentage GDP growth and inflation for the sample period from the first quarter of 1990 to the first quarter of 2009. For all countries, inflation largely mirrors GDP growth throughout the sample period. Industrialized economies are correlated, especially during times of crises. During the oil crisis in the early 1990s and the dot-com bubble at the beginning of the century nearly all countries' GDP growth rates were severely suppressed, while their inflation increased significantly. Note also that during the 1997 Asian financial crisis all the countries included here exhibited strong economic performance with high GDP growth and low inflation.

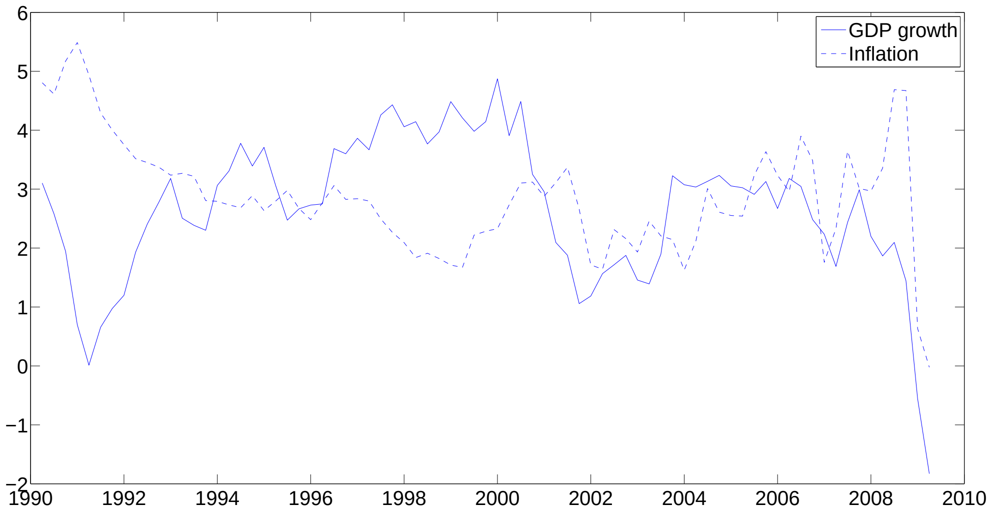


Figure 3.8: Annual % GDP growth and inflation, United States

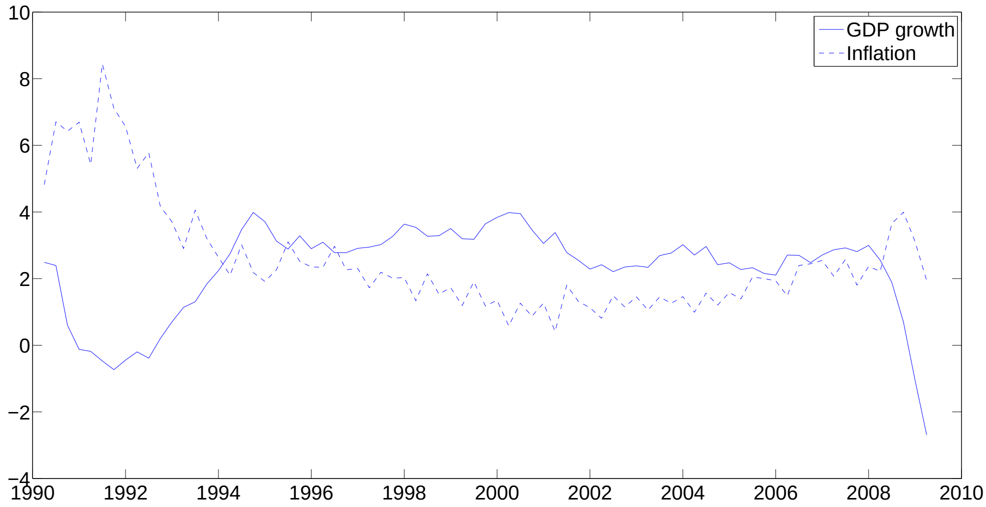


Figure 3.9: Annual % GDP growth and inflation, United Kingdom

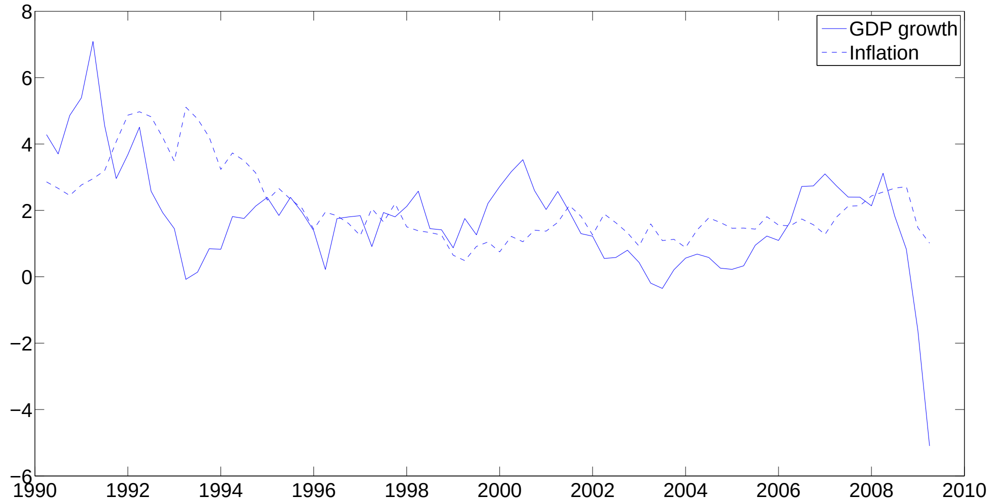


Figure 3.10: Annual % GDP growth and inflation, Germany

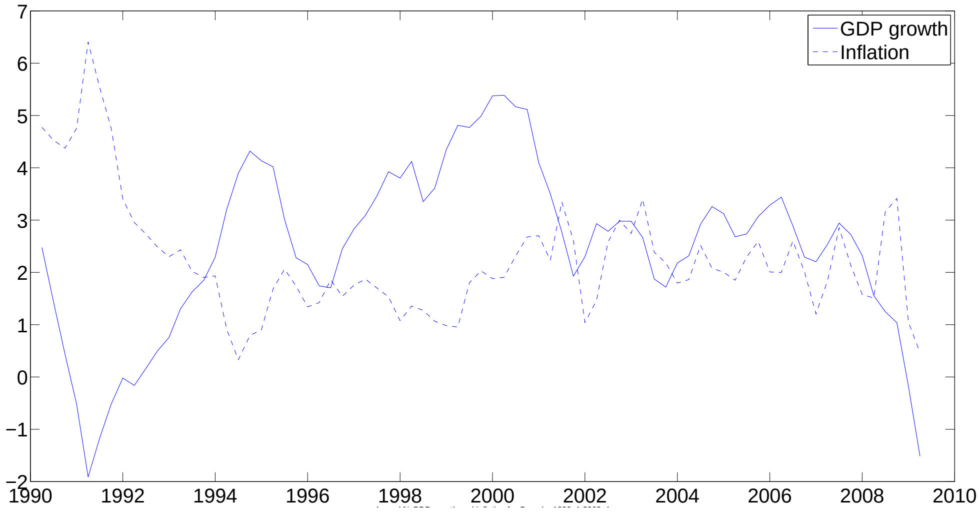


Figure 3.11: Annual % GDP growth and inflation, Canada

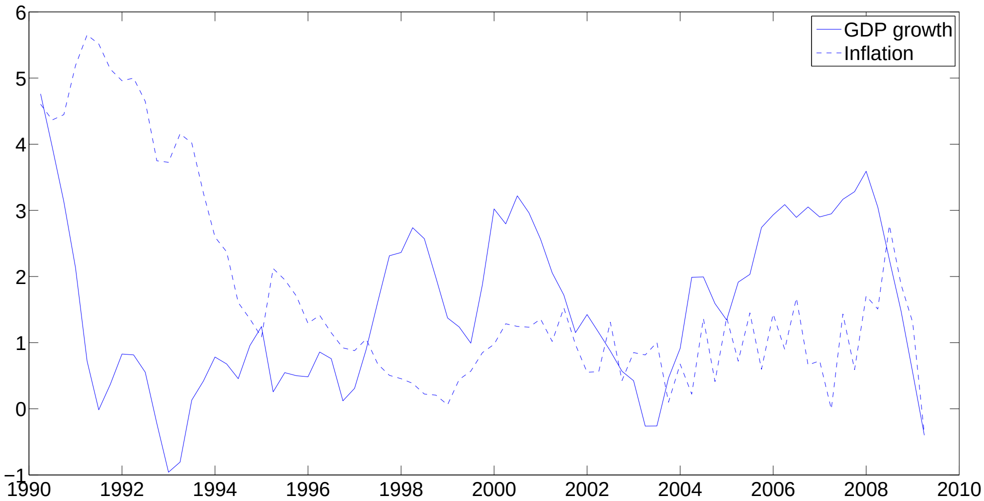


Figure 3.12: Annual % GDP growth and inflation, Switzerland

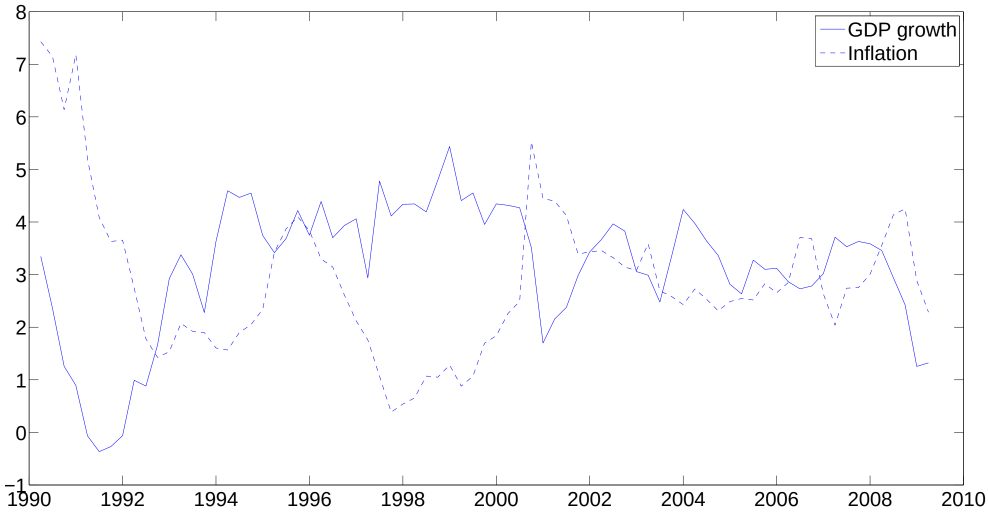


Figure 3.13: Annual % GDP growth and inflation, Australia

3.4 A term structure model with unspanned macro risk factors

This section introduces a Gaussian affine term structure model with unspanned macro risk factors that is able to price all bond yields consistently. The model follows the affine framework with unspanned macro risks proposed by Joslin, Priebsch and Singleton (2014).

The features of this framework are as follows. First, bond yields are driven by a small number of risk factors. Second, macro factors are unspanned by bond yield factors, that is, macro factors cannot be expressed as a linear combination of bond yields. Hence macro factors do not enter into the bond pricing equations under the risk-neutral measure $\mathbb{Q}$ to affect the current period bond yields. However, macro factors can be correlated with yield factors, and hence help to predict future yields. Therefore, the third feature of the model is that macro factors and yield factors follow a joint VAR without any restrictions under the physical measure $\mathbb{P}$.

3.4.1 The economy with unspanned macro risks

It is assumed that the economy at time t is driven by a state vector X_t, which includes all the risk factors in the economy. X_t is set to be an $N \times 1$ vector, meaning that there are in total N risk factors in the economy. Moreover, it follows a VAR(1) process under both the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$:

$$X_t = K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}}, \tag{3.1}$$

$$X_t = K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, \tag{3.2}$$

where $K_{0X}^{\mathbb{P}}$ and $K_{0X}^{\mathbb{Q}}$ are $N \times 1$ vectors, $K_{1X}^{\mathbb{P}}$, $K_{1X}^{\mathbb{Q}}$ and Σ_X are $N \times N$ matrices and both $\epsilon_t^{\mathbb{P}}$ and $\epsilon_t^{\mathbb{Q}}$ are $N \times 1$ vectors of independent shocks to various risk factors affecting the economy with mean zero and unit variance.

Specifically, X is made up of six risk factors. The first three are factors specifically affecting bond yields, which are set to be the first three principal components of bond yields. The other three are macro factors, which include the first difference of Asian reserves, a country's GDP growth and its inflation.

I denote the group of yield factors together by $\mathcal{P}_t$, and the group of macro factors by $\mathcal{M}_t$. Hence, the above process can be written in the following block structure.

Under the physical measure $\mathbb{P}$, equation (3.1) can be written as

$$\begin{bmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{bmatrix} = \begin{bmatrix} K_{0\mathcal{P}}^{\mathbb{P}} \\ K_{0\mathcal{M}}^{\mathbb{P}} \end{bmatrix} + \begin{bmatrix} K_{\mathcal{P}\mathcal{P}}^{\mathbb{P}} & K_{\mathcal{P}\mathcal{M}}^{\mathbb{P}} \\ K_{\mathcal{M}\mathcal{P}}^{\mathbb{P}} & K_{\mathcal{M}\mathcal{M}}^{\mathbb{P}} \end{bmatrix} \begin{bmatrix} \mathcal{P}_{t-1} \\ \mathcal{M}_{t-1} \end{bmatrix} + \Sigma_X \epsilon_t^{\mathbb{P}}.$$

Under the risk-neutral measure $\mathbb{Q}$, equation (3.2) can be written as

$$\begin{bmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{bmatrix} = \begin{bmatrix} K^{\mathbb{Q}}_{0\mathcal{P}} \\ K^{\mathbb{Q}}_{0\mathcal{M}} \end{bmatrix} + \begin{bmatrix} K^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} & 0 \\ K^{\mathbb{Q}}_{\mathcal{M}\mathcal{P}} & K^{\mathbb{Q}}_{\mathcal{M}\mathcal{M}} \end{bmatrix} \begin{bmatrix} \mathcal{P}_{t-1} \\ \mathcal{M}_{t-1} \end{bmatrix} + \Sigma_X \epsilon^{\mathbb{Q}}_t.$$

Note that while all blocks in X_t's $\mathbb{P}$ process are without restrictions, i.e. X_t follows an unrestricted VAR under $\mathbb{P}$, the upper right block of $K^{\mathbb{Q}}_{1X}$ is set to zero to take into account the fact that macro risk factors are unspanned by bond yield risk factors.

It is assumed that r_t, the one-period interest rate, is an affine function of yield factors $\mathcal{P}_t$ rather than X_t, i.e.

$$r_t = \rho_{0\mathcal{P}} + \rho'_{1\mathcal{P}} \mathcal{P}_t, \tag{3.3}$$

where $\rho_{0\mathcal{P}}$ is a scalar and $\rho_{1\mathcal{P}}$ is an $R \times 1$ vector, where R is the number of yield factors, which is equal to three with the current setup.

And the time-varying market price of risk vector that applies to the bond market $\lambda_{\mathcal{P}t}$ is also affine in $\mathcal{P}_t$:

$$\lambda_{\mathcal{P}t} = \lambda_{0\mathcal{P}} + \lambda_{1\mathcal{P}} \mathcal{P}_t. \tag{3.4}$$

Here $\lambda_{\mathcal{P}t}$ is an $R \times 1$ vector, $\lambda_{0\mathcal{P}}$ is an $R \times 1$ vector and $\lambda_{1\mathcal{P}}$ is an $R \times R$ matrix.

Hence, we obtain the bond-market-specific pricing kernel or the stochastic discount factor (SDF) $M_{\mathcal{P},t+1}$ of the bond market as

$$M_{\mathcal{P},t+1} = \exp(-r_t - \tfrac{1}{2}\lambda'_{\mathcal{P}t}\lambda_{\mathcal{P}t} - \lambda'_{\mathcal{P}t}\epsilon_{\mathcal{P},t+1}), \tag{3.5}$$

which can be used to consistently price all fixed-income assets. That is, we have the Euler equation

$$1 = E_t[M_{t+1}R_{t+1}], \tag{3.6}$$

where R_{t+1} is the one-period return on any asset in the bond market.

3.4.2 Zero-coupon bonds

To price nominal bonds, let P^{b}_{nt} denote the time-t price of the n-period nominal zero-coupon bond.

By the Euler equation, we have

$$P^{\mathrm{b}}_{nt} = E_t[M_{\mathcal{P},t+1}P^{\mathrm{b}}_{n-1,t+1}]. \tag{3.7}$$

Assuming the bond price is exponential affine in the bond yields' risk factors

$$P^{\mathrm{b}}_{nt} = \exp(A^{\mathrm{b}}_n + B^{\mathrm{b}\prime}_n \mathcal{P}_t), \tag{3.8}$$

where A^{b}_n is a scalar and B^{b}_n is an $R \times 1$ vector of bond price loadings on the bond risk factors, A^{b}_n and B^{b}_n can be solved by a set of Riccati equations with the boundary condition $P^{\mathrm{b}}_{0t} = 1$.

More specifically, the loadings are solved recursively as follows:

$$A^{\mathrm{b}}_n = -\rho_{0\mathcal{P}} + A^{\mathrm{b}}_{n-1} + {B^{\mathrm{b}}_{n-1}}' K^{\mathbb{Q}}_{0\mathcal{P}} + \tfrac{1}{2}{B^{\mathrm{b}}_{n-1}}'\Sigma_{\mathcal{P}}\Sigma'_{\mathcal{P}}B^{\mathrm{b}}_{n-1}, \tag{3.9}$$

$$B^{\mathrm{b}\prime}_n = -\rho'_{1\mathcal{P}} + {B^{\mathrm{b}}_{n-1}}' K^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} \tag{3.10}$$

with the starting values being $A^{\mathrm{b}}_n = 0$ and $B^{\mathrm{b}}_n = 0$.

And the bond yield at maturity n can be expressed as

$$y^{\mathrm{b}}_{nt} = -\frac{1}{n}\ln P^{\mathrm{b}}_{nt} = A_n + B'_n\mathcal{P}_t, \tag{3.11}$$

where $A_n = -A_n^{\mathrm{b}}/n$ and $B_n = -B_n^{\mathrm{b}}/n$.

3.5 Estimation

Section 3.5.1 summarizes the estimation strategy. Sections 3.5.2–3.5.7 present the estimation results for each country. In the latter, to investigate how macro factors, especially Asian reserve holdings, affect a country's yield curve, I look at the impulse responses of different yields to shocks to reserve holdings. I also investigate how shocks to different state factors contribute to yields' forecast variances using variance decompositions.

3.5.1 Estimation strategy

The estimation strategy follows Joslin, Priebsch and Singleton (2014). For each country, the state vector X_t comprises three yield factors, which are the first three principal components of zero-coupon bond yields of the fourteen maturities mentioned in the data section, and three macro factors, which are the first difference of Asian reserve holdings, the GDP growth and the inflation of the bond-issuing country. As the state vector X_t is entirely observable, the maximum likelihood estimates of the $\mathbb{P}$ parameters, i.e. $K_{0X}^{\mathbb{P}}$ and $K_{1X}^{\mathbb{P}}$, can be obtained from the VAR of X_t.

For the $\mathbb{Q}$ parameters, because the yield factors are chosen to be the principal components of bond yields, without loss of generality, yield factors can be rotated into a set of latent factors $\mathcal{L}_t$, whose $\mathbb{Q}$ process

$$\mathcal{L}_t = K_{0\mathcal{L}}^{\mathbb{Q}} + K_{\mathcal{L}\mathcal{L}}^{\mathbb{Q}} \mathcal{L}_{t-1} + \Sigma_{\mathcal{L}} \epsilon_{\mathcal{L}t}^{\mathbb{Q}} \tag{3.12}$$

can be written in its canonical form, as follows:

$$\mathcal{L}_t = \begin{bmatrix} \lambda_1^{\mathbb{Q}} & 0 & 0 \\ 0 & \lambda_2^{\mathbb{Q}} & 0 \\ 0 & 0 & \lambda_3^{\mathbb{Q}} \end{bmatrix} \mathcal{L}_{t-1} + \Sigma_{\mathcal{L}} \epsilon_{\mathcal{L}t}^{\mathbb{Q}},$$

i.e. $K_{0\mathcal{L}}^{\mathbb{Q}}$ is normalized to a zero vector. $K_{\mathcal{L}\mathcal{L}}^{\mathbb{Q}}$ is a diagonal matrix with real and distinct eigenvalues $\{\lambda_i^{\mathbb{Q}}\}$ on the diagonal, ordered by decreasing magnitude. $\Sigma_{\mathcal{L}}$ is the lower triangular Cholesky decomposition of $\Sigma_{\mathcal{L}}\Sigma_{\mathcal{L}}'$.

And the short rate under $\mathbb{Q}$ can be normalized such that

$$r_t = r_{\infty}^{\mathbb{Q}} + \iota' \mathcal{L}_t, \tag{3.13}$$

where ι is a vector of ones and $r_{\infty}^{\mathbb{Q}}$ can be interpreted as the risk-neutral long-run mean of the short rate.

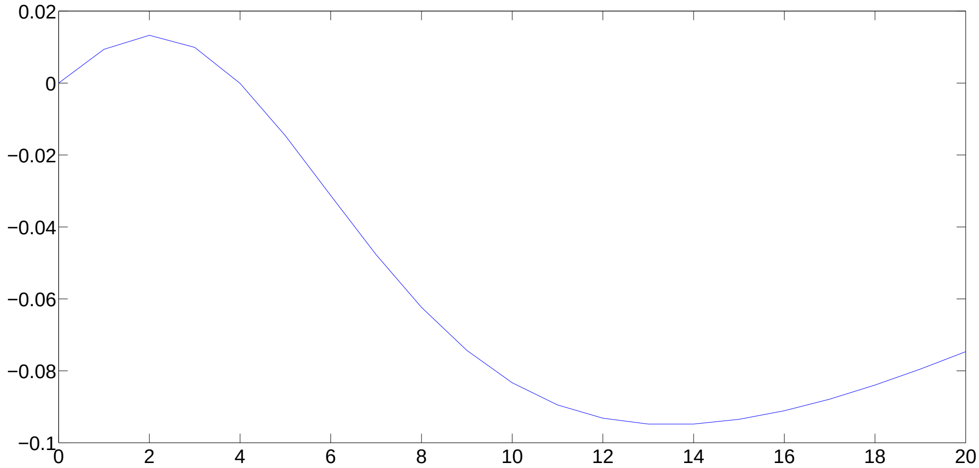
The set of $\mathbb{Q}$ parameters that needs to be estimated by the maximum likelihood estimation is reduced from ($\rho_{0\mathcal{P}}$, $\rho_{1\mathcal{P}}$, $K_{0\mathcal{P}}^{\mathbb{Q}}$, $K_{1\mathcal{P}}^{\mathbb{Q}}$, $\Sigma_{\mathcal{P}}$) to ($r_{\infty}^{\mathbb{Q}}$, $\{\lambda_i^{\mathbb{Q}}\}$ and $\Sigma_{\mathcal{P}}$).

3.5.2 Estimation results, United States

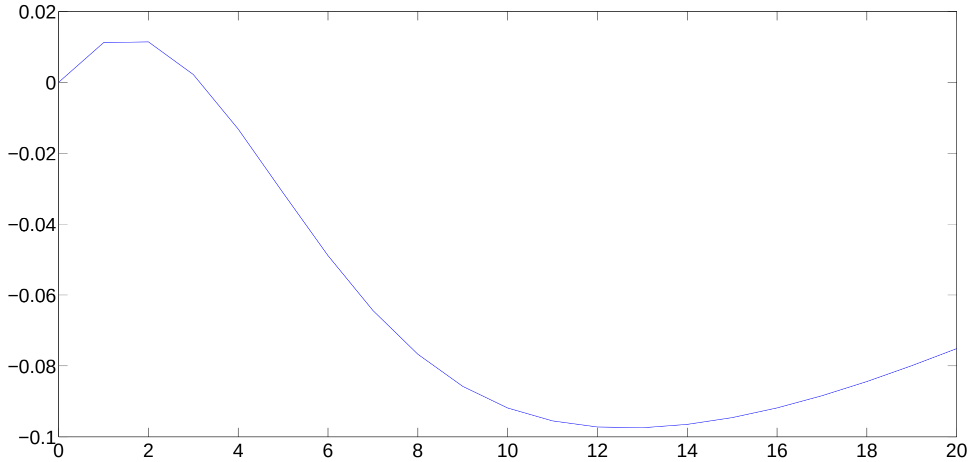
Impulse responses

Although macro factors are modeled as unspanned risk factors and do not enter into the pricing of contemporaneous bond yields, they co-evolve with yield factors in a VAR under the physical measure. Hence, we can trace out the dynamic responses of the yield curve from shocks to the macro variables by looking at the impulse responses of yields to macro shocks implied from the VAR. Figure 3.14 shows the impulse responses of yields of maturities of one, four and twenty quarters to shocks in reserves.

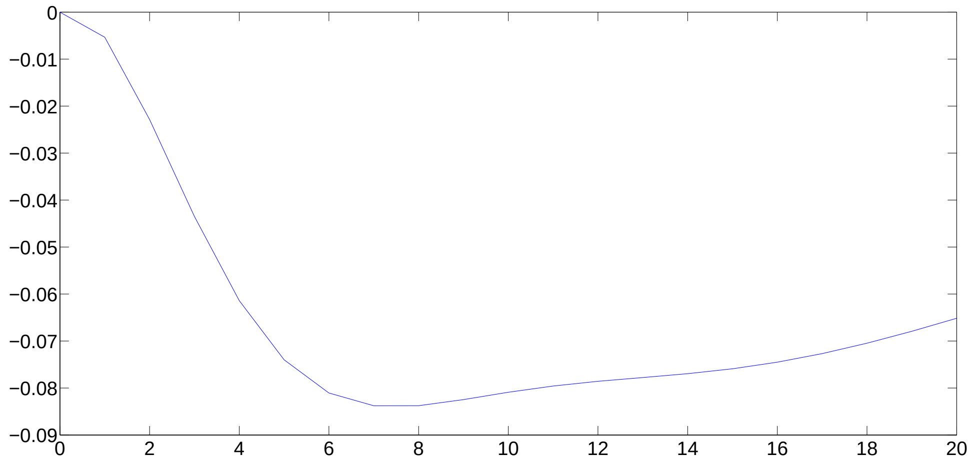
The detailed derivations of impulse response functions can be found in Appendix 2. From the figure, we can see that positive shocks to Asian countries' reserve holdings lower the U.S. bond yields of all maturities and across all horizons. In general, a one standard deviation shock to reserves (U.S.$45.8 billion) can lower bond yields of the three maturities by up to 10bp. While a positive shock to Asian reserves initially increases the U.S. bond yields of maturity of one quarter and four quarters for the first two quarters, then begins to have negative effects on yields, an increase in reserves has a pure negative effect on bond yield of 20 quarters. Moreover, as we move along the yield curve from its short end to its long end, a reserves shock reaches its maximum effect in less time. For example, while the response of the one-quarter yield to a shock to reserves reaches its peak at the 14th quarter, the response of the twenty-quarter yield reaches its maximum at the 7th quarter after the initial shock.



(a)



(b)



(c)

Figure 3.14: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), United States

Variance decompositions

This section uses variance decomposition techniques to investigate the proportion of the yields' forecast variance attributable to each state factor. Derivation for the variance decompositions are presented in Appendix 2. Table 3.4 lists the contribution of each factor to the h-step-ahead forecast of the one-quarter, four-quarter and twenty-quarter yields. By construction, we can read the table from two different dimensions. By fixing our attention at a particular p-quarter yield and varying the horizon, we can see how various factors (both yield factors and macro factors) affect the yield at different horizons. On the other hand, we can fix a particular forecast horizon and look at whether a particular factor has any effect across different maturities. The top row of the table lists the one-quarter-ahead forecast variance of the one-quarter yield explained by the three yield factors, i.e. the level, slope and curvature factor can explain 64.45%, 27.52% and 8.03% of the one-quarter-ahead forecast variance, respectively. Macro variables do not have explanatory power at the one-quarter-ahead forecast horizon for bond yields because they are unspanned and do not enter into the pricing of contemporaneous yields. But they do contribute to the yields' forecast variances at longer horizons. For example, for the yield at a maturity of one quarter, 1.98% of the twenty-quarter-ahead forecast variance is explained by reserves, 24.62% by GDP and 6.71% by inflation.

In order to see more clearly the explanatory power of macro factors, Table 3.5 lists the contribution of the macro factors to the h-step-ahead forecast variance of the short end, middle and long end of the yield curve. It shows that macro factors altogether explain a non-negligible amount of the variation in bond yields. At the short end of the yield curve, macro factors can explain 6.32% of the four-quarter-ahead forecast variance, and

Table 3.4: Variance decompositions, United States

	h	Yield factors			Macro factors		
		Level	Slope	Curvature	Reserves	GDP growth	Inflation
1	1	0.6445	0.2752	0.0803	0	0	0
quarter	4	0.8333	0.0740	0.0295	0.0003	0.0456	0.0172
yield	20	0.5516	0.0853	0.0300	0.0198	0.2462	0.0671
4	1	0.9483	0.0483	0.0033	0	0	0
quarter	4	0.9076	0.0152	0.0209	0.0002	0.0389	0.0171
yield	20	0.6069	0.0748	0.0220	0.0213	0.2150	0.0601
20	1	0.9651	0.0346	0.0003	0	0	0
quarter	4	0.9293	0.0408	0.0010	0.0023	0.0107	0.0159
yield	20	0.7190	0.0966	0.0036	0.0321	0.1129	0.0358

The table lists the contribution of factor i to the h-step-ahead forecast of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

Table 3.5: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, United States

	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	6.32%	33.3%	0.03%	1.98%
Middle (4 quarter yield)	5.63%	29.64%	0.02%	2.13%
Long end (20 quarter yield)	2.89%	18.08%	0.23%	3.21%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

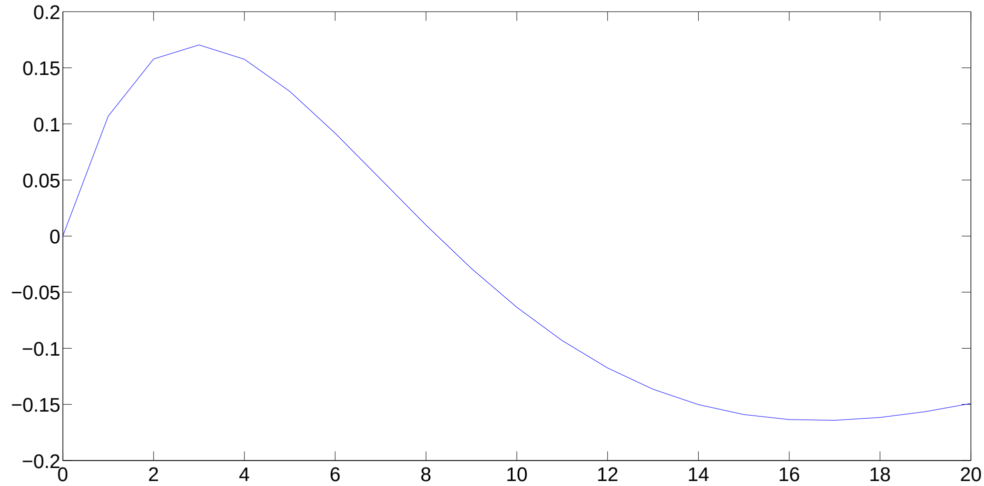
33.3% of the twenty-quarter-ahead forecast variance. As we move on to the middle and long end of the yield curve, the overall explanatory power of the macro factors starts to weaken. For bond yield at a maturity of twenty quarters, three macro factors together only explain 2.89% of the four-quarter-ahead forecast variance and 18.08% of the twenty-quarter-ahead forecast variance. However, for Asian reserves, contrary to the pattern observed for macro factors as a group, their ability in explaining forecast variance of bond yields increases with maturity. For example, at the short end of the yield curve, "reserves" is only able to explain

1.98% of the twenty-quarter-ahead forecast variance, but this increases to 3.21% for the long yield. Hence, we can see that reserves can have important implications for bond yields' forecast variance, especially at the long horizon. And their ability to explain yield variations increases with maturity, unlike the other macro factors namely GDP growth and inflation. Therefore, it is reasonable to say that while conventional macro factors, such as GDP growth and inflation, still explain a major proportion of the yield variations across all maturities and at all horizons, Asian reserves take an increasingly important role in explaining movements in the long end of the yield curve and at long forecast horizons. Hence, it is an important factor that affects the U.S. yield curve.

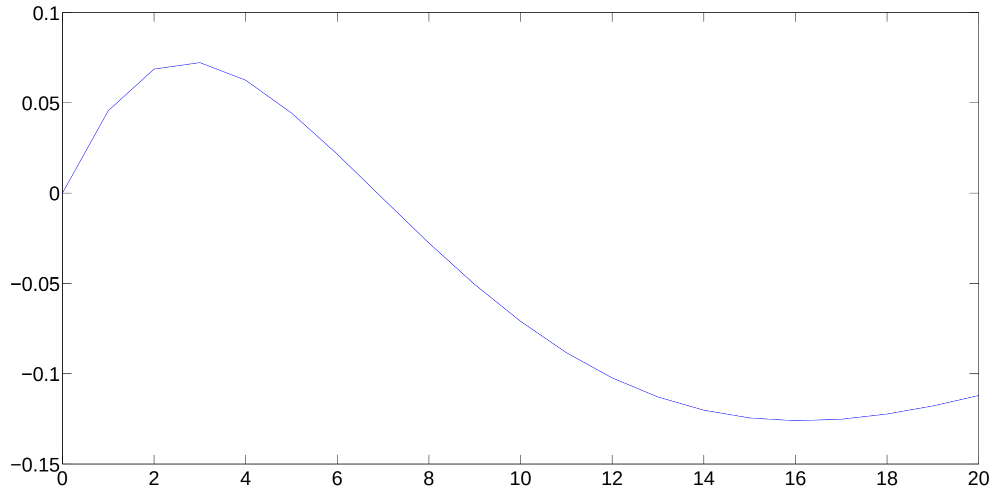
3.5.3 Estimation results, United Kingdom

Impulse responses

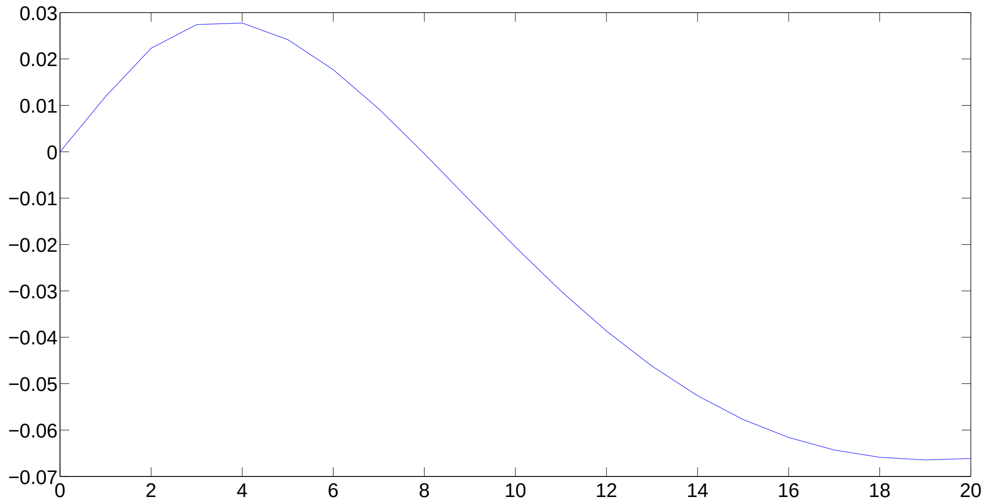
For the United Kingdom, impulse responses show that, across all maturities, a positive shock to Asian reserves initially increase U.K. bond yields for up to two years after the shock, then start to lower the yields, with the strongest effect at around year 4. Unlike the United States, whose yields at three maturities respond to reserves of comparable magnitudes, the magnitude of U.K. yields' responses quickly diminishes with maturity. While for the one-quarter yield, a positive shock to reserves can reduce the yield by more than 15bp, reserves can only reduce the twenty-quarter yield by less than 7bp.



(a)



(b)



(c)

Figure 3.15: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), United Kingdom

Variance decompositions

The contribution of macro factors toward yields' forecast variances for the United Kingdom follows a similar pattern to that of the United States, i.e. macro factors as a group have more explanatory power at the short end of the yield curve, while their contribution to yields' forecast variance is greater at the long forecast horizon and smaller at the short forecast horizon. When looking at reserves alone, their behavior follows the pattern described above for the macro factors group. It is worth noting that while for the United States shocks to Asian reserves affect the long end of the yield curve more than the short end in terms of contribution to the forecast variance in absolute values and as a proportion of the macro factors, it is the opposite for the United Kingdom. For the United Kingdom, reserves' explanatory power is strongest at the short end of the yield curve. At the four-quarter-ahead forecast horizon, reserves can explain 3.37% of the one-quarter yield's variance, but only 0.14% for the twenty-quarter yield. The opposing patterns of the United Kingdom and the United States can perhaps be understood by looking at the two countries' yield curves. Whereas the United States' yield curve is upward sloping, the United Kingdom's yield curve is relatively flat, with its one-quarter yield higher than the yields of all other maturities. Hence whereas investment of reserves in dollar-denominated assets is likely to be invested in U.S. long-term bonds when chasing higher yields, the same motive when applied to sterling-denominated assets will drive investments into short-term bonds rather than long-term bonds.

Table 3.6: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, United Kingdom

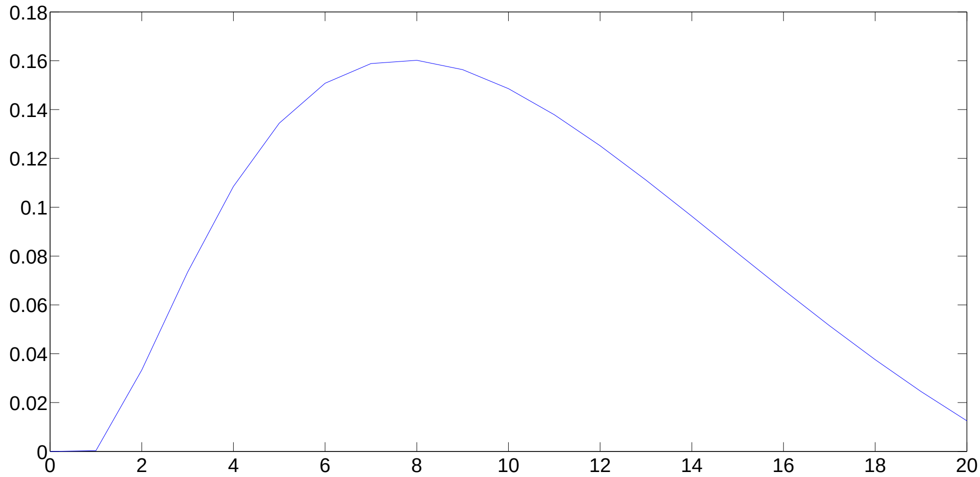
	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	5.84%	26.06%	3.37%	3.84%
Middle (4 quarter yield)	3.61%	24.04%	0.76%	2.58%
Long end (20 quarter yield)	2.26%	17.89%	0.14%	0.80%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

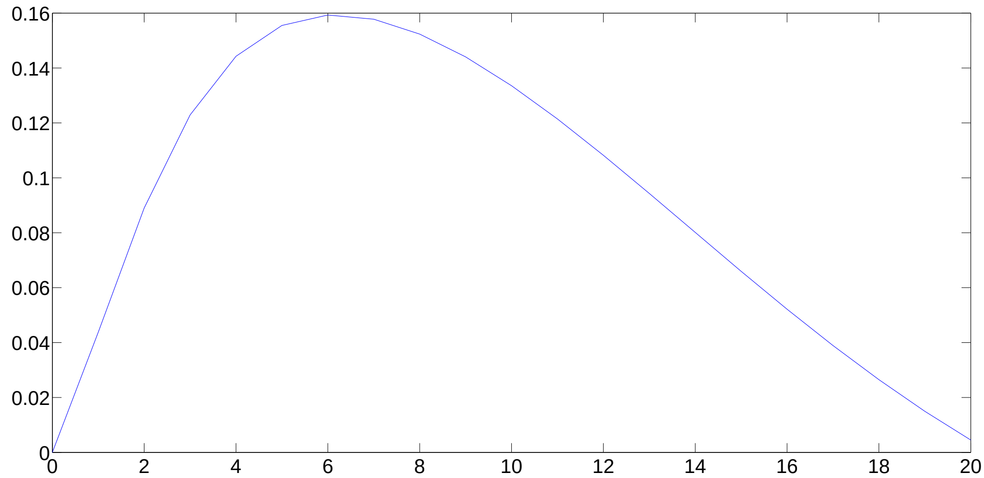
3.5.4 Estimation results, Germany

Impulse responses

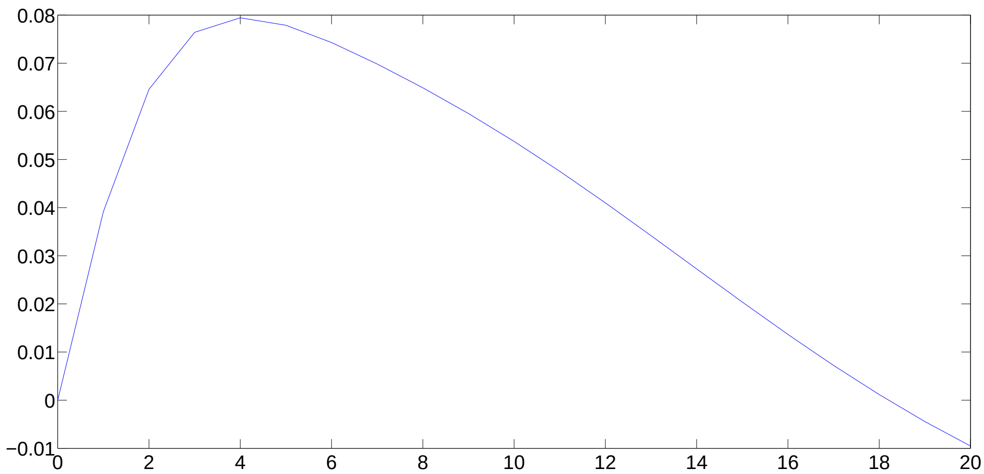
For Germany, although its average yields at all maturities closely track those of the U.S., their dynamic responses to reserves shocks are very different. German bond yields largely react positively to reserves shocks across all maturities. Yields of longer maturities react less to shocks to reserves. Also the reserves' effect on bond yields is strongest at around four to eight quarters after the initial shock.



(a)



(b)



(c)

Figure 3.16: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), Germany

Variance decompositions

For Germany, reserves' explanatory power is strongest for mid-maturity bonds. For a yield of a maturity of four quarters, reserves can explain 2.45% of the forecast variance at the four-quarter forecast horizon. And at the twenty-quarter forecast horizon, reserves' explanatory power increases to nearly 4%.

Table 3.7: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, Germany

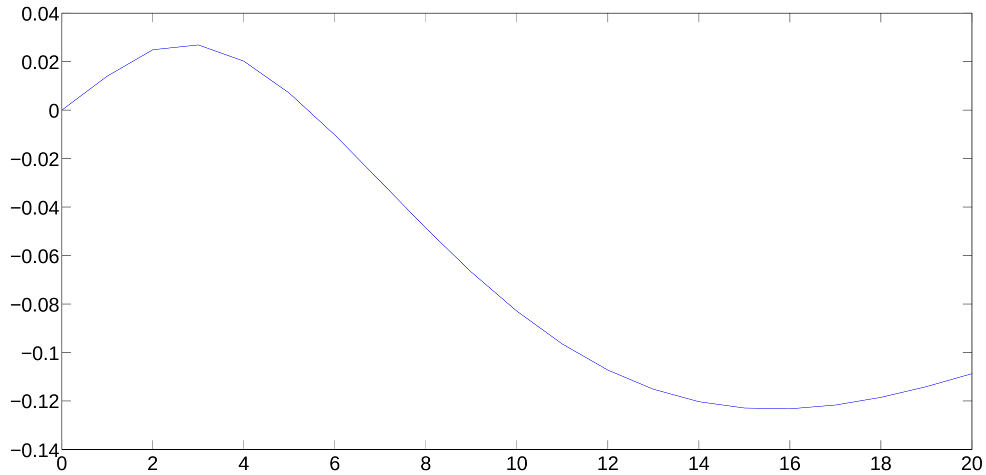
	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	24.31%	52.43%	0.63%	3.37%
Middle (4 quarter yield)	30.17%	52.61%	2.45%	3.97%
Long end (20 quarter yield)	21.41%	39.68%	1.84%	1.68%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

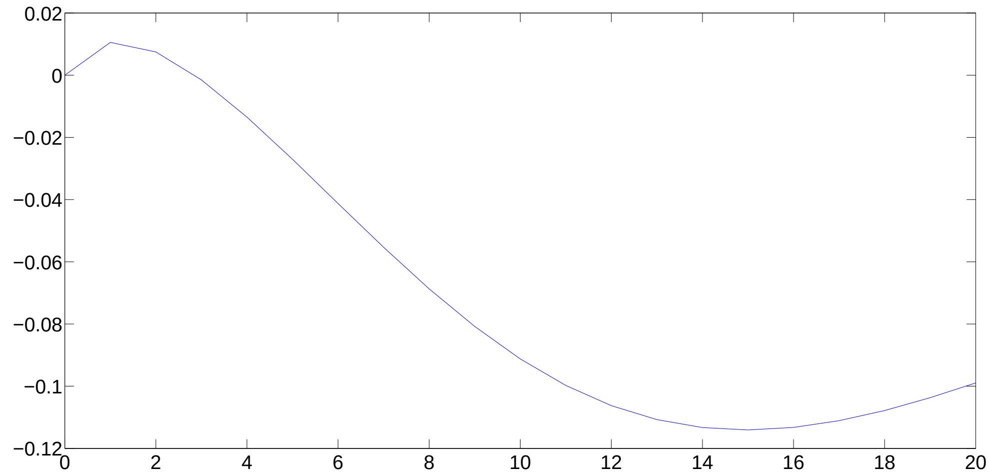
3.5.5 Estimation results, Canada

Impulse responses

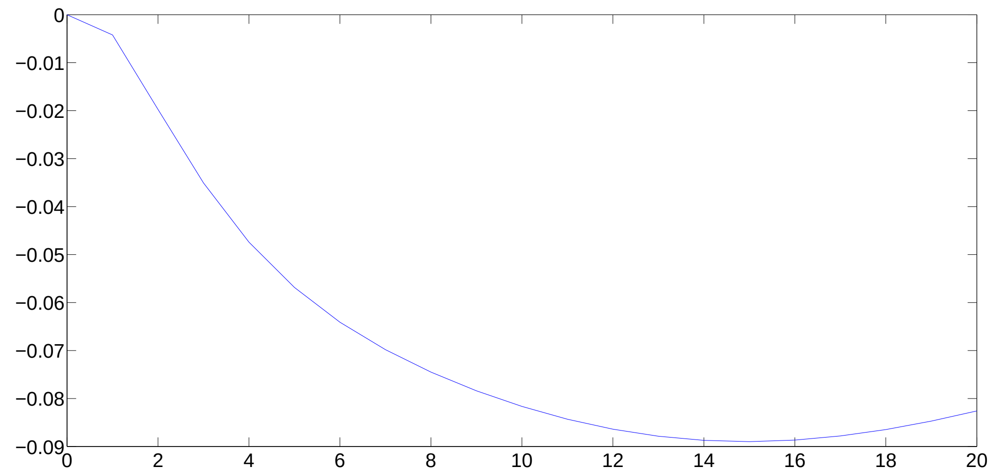
Canadian yields' impulse responses to reserves holdings shocks are very similar to those of the United States. For both countries, yields respond at similar magnitudes, with shocks to reserves having an effect at slightly longer horizons for Canada.



(a)



(b)



(c)

Figure 3.17: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), Canada

Variance decompositions

For Canada, the percentages of yields' forecast variances explained by reserves is largest for the long maturity bond and at the long forecast horizon, just as observed for the United States. Another similarity is that, while the explanatory power of reserves increases with maturity, the opposite is true for the explanatory power of all macro factors as a group. For example, at the twenty-quarter-ahead forecast horizon, the contribution of reserves shocks towards yields forecast variance increases from 2.65% for the one-quarter yield to 3.48% for the twenty-quarter yield. In contrast, shocks to macro factors altogether contribute 23.24% of the forecast variance of the one-quarter yield and only 16.78% of the variance of the twenty-quarter yield.

Table 3.8: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, Canada

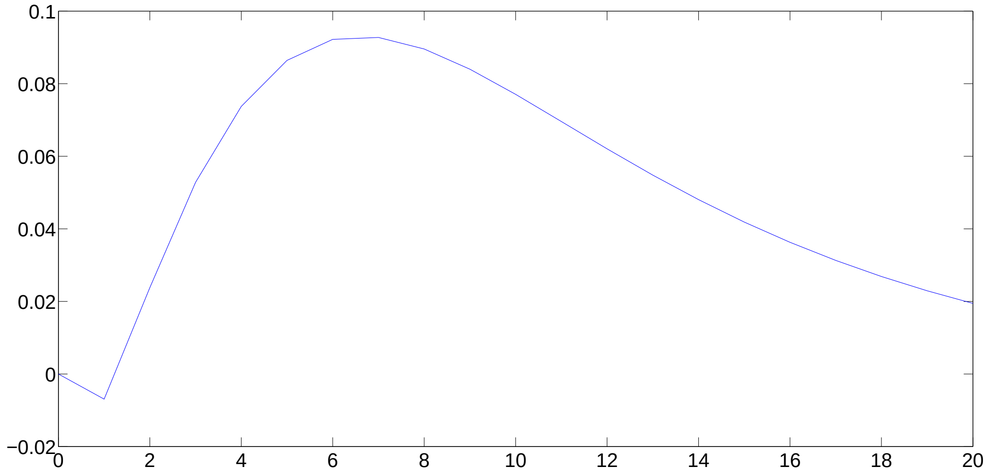
	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	3.75%	23.24%	0.09%	2.65%
Middle (4 quarter yield)	4.94%	22.06%	0.01%	3.00%
Long end (20 quarter yield)	4.66%	16.78%	0.20%	3.48%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

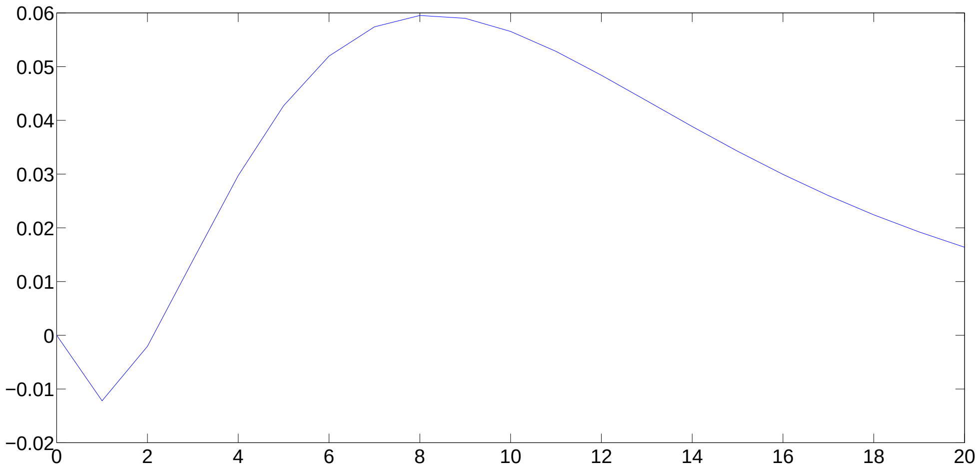
3.5.6 Estimation results, Switzerland

Impulse responses

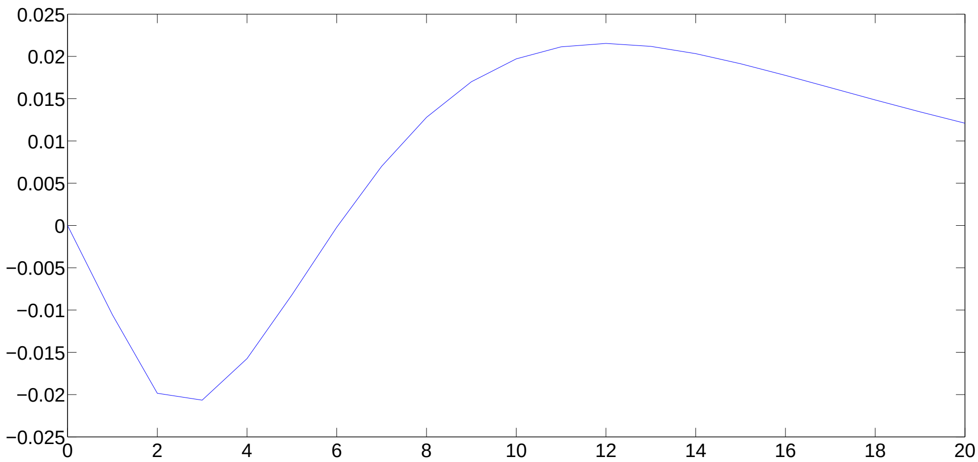
Like German yields, Swiss yields generally increase with an increase in reserves. Swiss yields initially have negative responses to reserves shocks, but the magnitudes are very small across maturities. And the reserves' effect is strongest between two and three years after the initial shock.



(a)



(b)



(c)

Figure 3.18: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), Switzerland

Variance decompositions

Looking at the yields' forecast variance decompositions, macro factors generally have low explanatory power for Swiss bond yields compared with yields' variance decompositions of other countries. For example, for Germany, shocks to macro factors account for 52.43% of the twenty-quarter-ahead forecast variance of the short maturity bond. For Switzerland, this number drops to 17.10%. As we move to longer maturities, the explanatory power of both macro factors and reserves decreases.

Table 3.9: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, Switzerland

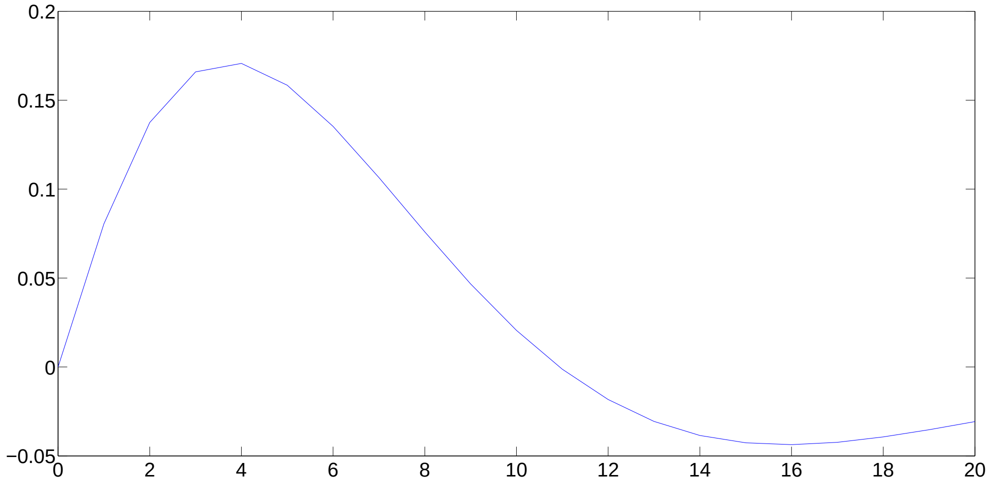
	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	10.59%	17.10%	0.29%	2.12%
Middle (4 quarter yield)	6.16%	14.77%	0.04%	1.25%
Long end (20 quarter yield)	1.59%	8.15%	0.20%	0.41%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

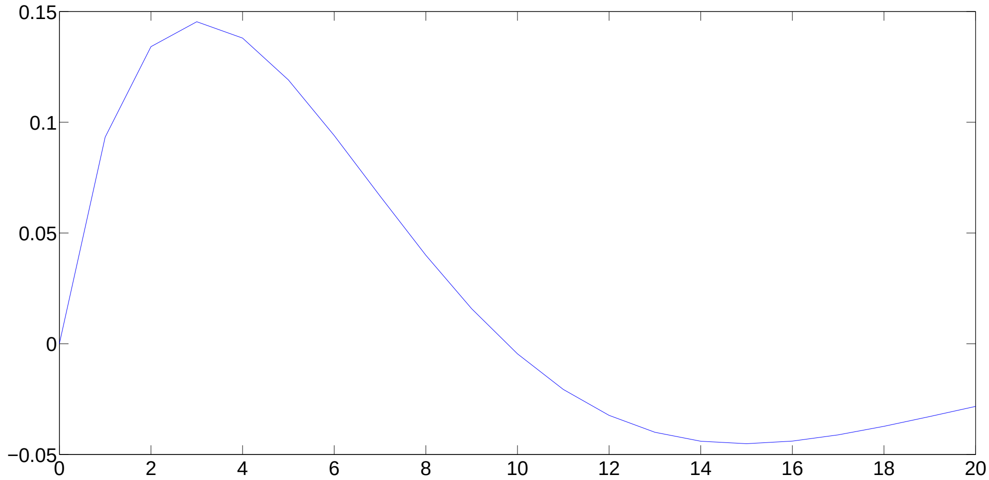
3.5.7 Estimation results, Australia

Impulse responses

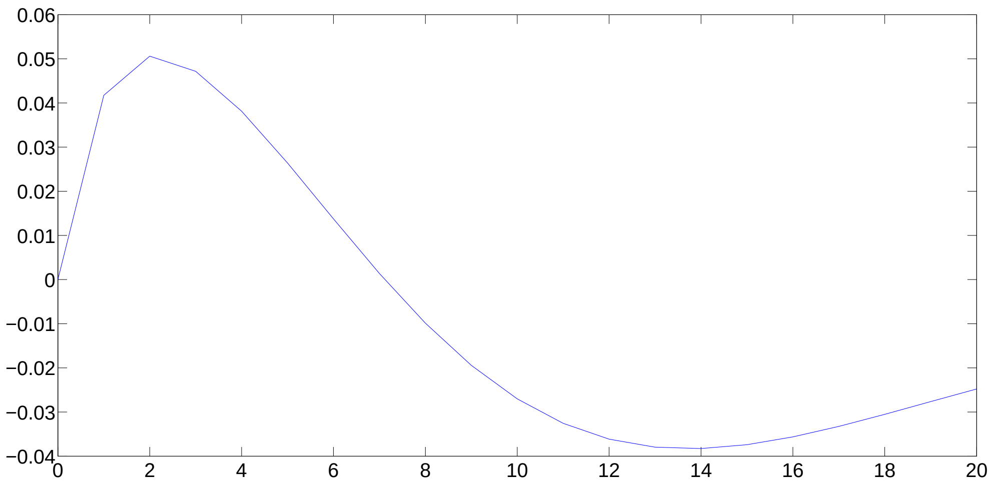
Looking at Australia, shocks to reserves initially increase bond yields. Then, after about two to three years, an initial increase in reserves lowers bond yields.



(a)



(b)



(c)

Figure 3.19: Impulse response functions of yields to a one standard deviation shock to Asian reserves. Impulse responses for yields of maturities of one quarter (top panel), four quarters (middle panel) and twenty quarters (bottom panel), Australia

Variance decompositions

The explanatory power of reserves or macro factors as a group is strongest at the short end of the yield curve. Whereas reserves explain only 0.59% and 0.99% at the four- and twenty-quarter forecast horizons for the long-maturity (twenty-quarter) yield, shocks to reserves contribute to 4.64% and 6.78%, respectively, of the four-quarter and twenty-quarter-ahead forecast variance of the short maturity (one-quarter) yield.

Table 3.10: Proportion of yields' forecast variance explained by macro factors (reserves, GDP growth, inflation) and reserves, Australia

	Macro factors		Reserves	
Forecast horizon	4 quarter	20 quarter	4 quarter	20 quarter
Short end (1 quarter yield)	8.99%	19.15%	4.64%	6.78%
Middle (4 quarter yield)	8.27%	15.68%	3.68%	4.65%
Long end (20 quarter yield)	2.64%	6.14%	0.59%	0.99%

The table lists the contribution of the macro factors (reserves, GDP growth, inflation) and the contribution of reserves to the h-step-ahead (four quarters, twenty quarters) forecast variance of the short end (one-quarter yield), middle (four-quarter yield) and long end (twenty-quarter yield) of the yield curve.

3.6 Conclusion

This chapter studies the effect of Asian countries' reserve holdings on the yield curves of six industrialized countries: the United States, the United Kingdom, Germany, Canada, Switzerland and Australia. A Gaussian affine term structure model with three yield factors and three unspanned macro factors is estimated to fit the yield curve of each country. Yields factors and macro factors are set to follow an unrestricted VAR under the physical measure. Hence, impulse responses and variance decompositions of yields to all factor shocks can be obtained. Impulse responses show that a one standard deviation shock to Asian reserve holdings can move bond yields of the above countries by up to 18bp during the first five years after the initial shock. And variance decompositions show that a significant proportion of the yields' forecast variance can be attributed to reserves. For the five-year-ahead forecast horizon, the highest proportions of variance explained by reserves can be as large as 6.78%. Moreover, the explanatory power of reserves generally increases with forecast horizon across all countries and maturities. Therefore, Asian countries' reserve holdings are an important factor affecting the international yield curves.

Bibliography

Ang, A. and M. Piazzesi, "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables* 1," *Journal of Monetary economics*, 2003, *50* (4), 745–787.

Ang, Andrew and Jun Liu, "How to discount cashflows with time-varying expected returns," *The Journal of Finance*, 2004, *59* (6), 2745–2783.

_ **and Maxim Ulrich**, "Nominal bonds, real bonds, and equity," *Real Bonds, and Equity (March 5, 2012)*, 2012.

Bansal, Ravi and Amir Yaron, "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 2004, *59* (4), 1481–1509.

Barro, Robert J, "Rare disasters, asset prices, and welfare costs," *American Economic Review*, 2009, *99* (1), 243–264.

Bauer, Michael D, Glenn D Rudebusch, and Jing Cynthia Wu, "Correcting estimation bias in dynamic term structure models," *Journal of Business & Economic Statistics*, 2012, *30* (3), 454–467.

Bekaert, Geert and Steven R Grenadier, "Stock and Bond Pricing in an Affine Economy," *EFA 2002 Berlin Meetings Working Paper*, 2001.

_ , Eric Engstrom, and Steven R Grenadier, "Stock and bond returns with moody investors," *Journal of Empirical Finance*, 2010, *17* (5), 867–894.

Beltran, Daniel O, Maxwell Kretchmer, Jaime Marquez, and Charles P Thomas, "Foreign holdings of US Treasuries and US Treasury yields," *Journal of International Money and Finance*, 2013, *32*, 1120–1143.

Bernanke, Ben, Vincent Reinhart, and Brian Sack, "Monetary policy alternatives at the zero bound: An empirical assessment," *Brookings papers on economic activity*, 2004, *2004* (2), 1–100.

Campbell, John Y and John H Cochrane, "By force of habit: A consumption-based explanation of aggregate stock market behavior," *The Journal of Political Economy*, 1999, *107* (2), 205–251.

Chernov, Mikhail and Philippe Mueller, "The term structure of inflation expectations," *Journal of financial economics*, 2012, *106* (2), 367–394.

Cochrane, John H and Monika Piazzesi, "Bond risk premia," *American Economic Review*, 2005, *95* (1), 138–160.

Dai, Qiang and Kenneth J Singleton, "Specification analysis of affine term structure models," *The Journal of Finance*, 2000, *55* (5), 1943–1978.

Duffee, Gregory R, "Information in (and not in) the term structure," *Review of Financial Studies*, 2011, *24* (9), 2895–2934.

Duffie, D. and R. Kan, "A yield-factor model of interest rates," *Mathematical finance*, 1996, *6* (4), 379–406.

Gabaix, Xavier, "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance," *The Quarterly Journal of Economics*, 2012, *127* (2), 645–700.

Goodfriend, Marvin, "Interest rates and the conduct of monetary policy," in "Carnegie-Rochester conference series on public policy," Vol. 34 Elsevier 1991, pp. 7–30.

Greenspan, Alan, "Testimony of Chairman Alan Greenspan," in "Federal Reserve Boards semiannual Monetary Policy Report to the Congress Before the Committee on Banking, Housing, and Urban Affairs, US Senate," Vol. 16 2005.

Jardet, Caroline, Alain Monfort, and Fulvio Pegoraro, "No-arbitrage near-cointegrated VAR (p) term structure models, term premia and GDP growth," *Journal of Banking & Finance*, 2013, *37* (2), 389–402.

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

_ , Kenneth J Singleton, and Haoxiang Zhu, "A new perspective on Gaussian dynamic term structure models," *Review of Financial Studies*, 2011, *24* (3), 926–970.

_ , Marcel Priebsch, and Kenneth J Singleton, "Risk premiums in dynamic term structure models with unspanned macro risks," *The Journal of Finance*, 2014, *69* (3), 1197–1233.

Koijen, Ralph SJ, Hanno N Lustig, and Stijn Van Nieuwerburgh, "The cross-section and time-series of stock and bond returns," 2013.

Lettau, Martin and Jessica A. Wachter, "Why is long-horizon equity less risky? A duration-based explanation of the value premium," *The Journal of Finance*, 2007, *62* (1).

_ and _ , "The term structures of equity and interest rates," *Journal of Financial Economics*, 2011, *101* (1), 90–113.

Litterman, R.B. and J. Scheinkman, "Common factors affecting bond returns," *The Journal of Fixed Income*, 1991, *1* (1), 54–61.

Rudebusch, Glenn D, Eric T Swanson, and Tao Wu, "The Bond Yield Conundrum from a Macro-Finance Perspective," *Monetary and Economic Studies (Special Edition)*, 2006.

Sierra, Jesus, "International Capital Flows and Bond Risk Premia," *The Quarterly Journal of Finance*, 2014, *4* (01), 1450001.

van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, "On the Timing and Pricing of Dividends," *American Economic Review*, 2012, *102* (4), 1596–1618.

Wachter, Jessica A, "A consumption-based model of the term structure of interest rates," *Journal of Financial economics*, 2006, *79* (2), 365–399.

Warnock, Francis E and Veronica Cacdac Warnock, "International capital flows and US interest rates," *Journal of International Money and Finance*, 2009, *28* (6), 903–919.

Wright, Jonathan H, "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset," *The American Economic Review*, 2011, *101* (4), 1514.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," *Journal of the American Statistical Association*, 1962, *57* (298), 348–368.

Appendix 1

A1.1 Solutions for asset prices

Let P^{d}_{nt} denote the time-t price of an asset that pays off the aggregate dividend at time $t+n$, i.e. the price of the zero-coupon equity of maturity n. Its one-period holding period return can be specified as the return we get from buying it at time t at the price P^{d}_{nt} and selling it at time $t+1$ at the price $P^{\mathrm{d}}_{n-1,t+1}$:

$$R^{\mathrm{d}}_{n,t+1} = \frac{P^{\mathrm{d}}_{n-1,t+1}}{P^{\mathrm{d}}_{nt}} = \frac{P^{\mathrm{d}}_{n-1,t+1}/D_{t+1}}{P^{\mathrm{d}}_{nt}/D_t}\frac{D_{t+1}}{D_t}.$$

The Euler equation can be written as

$$\frac{P^{\mathrm{d}}_{nt}}{D_t} = E_t\left[M_{t+1}\frac{P^{\mathrm{d}}_{n-1,t+1}}{D_{t+1}}\frac{D_{t+1}}{D_t}\right]$$

with boundary condition

$$\frac{P^{\mathrm{d}}_{0t}}{D_t} = 1.$$

Assume the solution to P^{d}_{nt}/D_t is exponential affine in the state vector

$$\frac{P^{\mathrm{d}}_{nt}}{D_t} = \exp(A^{\mathrm{d}}_n + B^{\mathrm{d}\prime}_n X_t).$$

Then, given the left-hand side of the Euler equation is exponential affine in X_t, we need to write the right-hand side of the Euler exponential affine in X_t as well.

Just like the above equation, we have

$$\begin{aligned}
\frac{P^{\mathrm{d}}_{n-1,t+1}}{D_{t+1}} &= \exp(A^{\mathrm{d}}_{n-1} + {B^{\mathrm{d}}_{n-1}}' X_{t+1}) \\
&= \exp(A^{\mathrm{d}}_{n-1} + {B^{\mathrm{d}}_{n-1}}'(K^{\mathbb{P}}_{0X} + (K^{\mathbb{P}}_{1X} + I)X_t + \Sigma_X \epsilon_{t+1})) \\
&= \exp(A^{\mathrm{d}}_{n-1} + {B^{\mathrm{d}}_{n-1}}' K^{\mathbb{P}}_{0X} + {B^{\mathrm{d}}_{n-1}}'(K^{\mathbb{P}}_{1X} + I)X_t + {B^{\mathrm{d}}_{n-1}}' \Sigma_X \epsilon_{t+1}).
\end{aligned}$$

Given dividend growth is exponential affine in the state variables, we have

$$\begin{aligned}
\frac{D_{t+1}}{D_t} &= \exp(\Delta d_{t+1}) = \exp(\delta_{0X} + \delta'_{1X} X_{t+1}) \\
&= \exp(\delta_{0X} + \delta'_{1X}(K^{\mathbb{P}}_{0X} + (K^{\mathbb{P}}_{1X} + I)X_t + \Sigma_X \epsilon_{t+1})) \\
&= \exp(\delta_{0X} + \delta'_{1X} K^{\mathbb{P}}_{0X} + \delta'_{1X}(K^{\mathbb{P}}_{1X} + I)X_t + \delta'_{1X} \Sigma_X \epsilon_{t+1}).
\end{aligned}$$

Finally, recall the SDF is exponential affine in the state variables

$$\begin{aligned}
M_{t+1} &= \exp(-r_t - \tfrac{1}{2}\lambda'_t \lambda_t - \lambda'_t \epsilon_{t+1}) \\
&= \exp(-\rho_{0X} - \rho'_{1X} X_t - \tfrac{1}{2}\lambda'_t \lambda_t - \lambda'_t \epsilon_{t+1}).
\end{aligned}$$

Substituting the above affine forms (in X_t) into the Euler equation, we get [2]

$$
\begin{aligned}
&\exp(A_n^{\mathrm{d}} + B_n^{\mathrm{d}\prime} X_t) \\
&\quad = E_t[\exp(-\rho_{0X} - \rho_{1X}' X_t - \tfrac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}) \\
&\qquad\qquad \times \exp(A_{n-1}^{\mathrm{d}} + B_{n-1}^{\mathrm{d}}{}' K_{0X}^{\mathbb{P}} + B_{n-1}^{\mathrm{d}}{}'(K_{1X}^{\mathbb{P}} + I)X_t + B_{n-1}^{\mathrm{d}}{}'\Sigma_X \epsilon_{t+1}) \\
&\qquad\qquad \times \exp(\delta_{0X} + \delta_{1X}' K_{0X}^{\mathbb{P}} + \delta_{1X}'(K_{1X}^{\mathbb{P}} + I)X_t + \delta_{1X}'\Sigma_X \epsilon_{t+1})] \\
&\quad = E_t[\exp(-\rho_{0X} + \delta_{0X} - \tfrac{1}{2}\lambda_t'\lambda_t + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})' K_{0X}^{\mathbb{P}} \\
&\qquad\qquad\quad + (-\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{1X}^{\mathbb{P}} + I))X_t \\
&\qquad\qquad\quad + (-\lambda_t' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X)\epsilon_{t+1})] \\
&\quad = \exp(-\rho_{0X} + \delta_{0X} - \tfrac{1}{2}\lambda_t'\lambda_t + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})' K_{0X}^{\mathbb{P}} \\
&\qquad\qquad + (-\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{1X}^{\mathbb{P}} + I))X_t \\
&\qquad\qquad + \tfrac{1}{2}\lambda_t'\lambda_t + \tfrac{1}{2}(\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\Sigma_X'(\delta_{1X} + B_{n-1}^{\mathrm{d}}) \\
&\qquad\qquad - (\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\lambda_t) \\
&\quad = \exp(-\rho_{0X} + \delta_{0X} + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})' K_{0X}^{\mathbb{P}} \\
&\qquad\qquad + (-\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{1X}^{\mathbb{P}} + I))X_t \\
&\qquad\qquad + \tfrac{1}{2}(\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\Sigma_X'(\delta_{1X} + B_{n-1}^{\mathrm{d}}) \\
&\qquad\qquad - (\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\lambda_{0X} - (\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\lambda_{1X}X_t) \\
&\quad = \exp(-\rho_{0X} + \delta_{0X} + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{0X}^{\mathbb{P}} - \Sigma_X\lambda_{0X}) \\
&\qquad\qquad + \tfrac{1}{2}(\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\Sigma_X'(\delta_{1X} + B_{n-1}^{\mathrm{d}}) \\
&\qquad\qquad + (-\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'((K_{1X}^{\mathbb{P}} + I) - \Sigma_X\lambda_{1X}))X_t) \\
&\quad = \exp(-(\rho_{0X} - \delta_{0X}) + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})' K_{0X}^{\mathbb{Q}} \\
&\qquad\qquad + \tfrac{1}{2}(\delta_{1X} + B_{n-1}^{\mathrm{d}})'\Sigma_X\Sigma_X'(\delta_{1X} + B_{n-1}^{\mathrm{d}}) \\
&\qquad\qquad + (-\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{1X}^{\mathbb{Q}} + I))X_t).
\end{aligned}
$$

[2] Note $\lambda_t = \lambda_0 + \lambda_1 X_t$ is also affine in X_t, but it is not decomposed here as, when taking the expectation of the SDF, the two terms containing λ_t cancel out, or the terms that are quadratic in X_t cancel out. Then, after the cancellation, we can decompose λ_t into $\lambda_0 + \lambda_1 X_t$ and get an expression affine in X_t.

Hence we have

$$\begin{aligned} A_n^{\mathrm{d}} &= -(\rho_{0X} - \delta_{0X}) + A_{n-1}^{\mathrm{d}} + (\delta_{1X} + B_{n-1}^{\mathrm{d}})' K_{0X}^{\mathbb{Q}} \\ &\qquad + \tfrac{1}{2}(\delta_{1X} + B_{n-1}^{\mathrm{d}})' \Sigma_X \Sigma_X' (\delta_{1X} + B_{n-1}^{\mathrm{d}}), \\ B_n^{\mathrm{d}\prime} &= -\rho_{1X}' + (\delta_{1X} + B_{n-1}^{\mathrm{d}})'(K_{1X}^{\mathbb{Q}} + I), \end{aligned}$$

with starting values $A_0^{\mathrm{d}} = 0$ and $B_0^{\mathrm{d}} = 0$.

A1.2 Proof of Proposition 1

Given

$$\begin{aligned} \Delta X_t &= K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, \\ r_t &= \rho_{0X} + \rho_{1X}{}' X_t, \\ \Delta d_t &= \delta_{0X} + \delta_{1X}{}' X_t, \end{aligned}$$

assume X_t is stationary under $\mathbb{Q}$ with unconditional mean μ such that $\mu = -(K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}$.

Letting $X_t^* = X_t - \mu = X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}$, the state process becomes

$$\begin{aligned} \Delta X_t^* &= K_{0X}^{\mathbb{Q}} + K_{1X}^{\mathbb{Q}} [X_{t-1}^* - (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}] + \Sigma_X \epsilon_t^{\mathbb{Q}} \\ &= K_{1X}^{\mathbb{Q}} X_{t-1}^* + \Sigma_X \epsilon_t^{\mathbb{Q}}. \end{aligned}$$

For ease of exposition, assume $K_{1X}^{\mathbb{Q}}$ has non-zero, real and distinct eigenvalues $\lambda^{\mathbb{Q}}$. Then, by eigendecomposition, we can write

$$K_{1X}^{\mathbb{Q}} = A^{\mathbb{Q}} \operatorname{diag}(\lambda^{\mathbb{Q}}) (A^{\mathbb{Q}})^{-1},$$

and hence

$$\Delta X_t^* = A^{\mathbb{Q}} \operatorname{diag}(\lambda^{\mathbb{Q}})(A^{\mathbb{Q}})^{-1} X_{t-1}^* + \Sigma_X \epsilon_t^{\mathbb{Q}}.$$

Letting $X_t^{**} = (A^{\mathbb{Q}})^{-1} X_t^*$, the state process becomes

$$\Delta X_t^{**} = \operatorname{diag}(\lambda^{\mathbb{Q}}) X_{t-1}^{**} + (A^{\mathbb{Q}})^{-1} \Sigma_X \epsilon_t^{\mathbb{Q}}.$$

Given $X_t^{**} = (A^{\mathbb{Q}})^{-1} X_t^*$ and $X_t^* = X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}$, we have $X_t^{**} = (A^{\mathbb{Q}})^{-1}[X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}]$, or $X_t = A^{\mathbb{Q}} X_t^{**} - (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}$. The short rate equation can then be written as

$$\begin{aligned}
r_t &= \rho_{0X} + {\rho_{1X}}' X_t \\
&= \rho_{0X} + {\rho_{1X}}'[A^{\mathbb{Q}} X_t^{**} - (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}] \\
&= r_\infty^{\mathbb{Q}} + \iota' X_t^{***},
\end{aligned}$$

where $r_\infty^{\mathbb{Q}} = \rho_{0X} - {\rho_{1X}}'(K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}$ and $X_t^{***} = \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}}) X_t^{**}$.

With $X_t^{***} = \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}}) X_t^{**}$, rewrite the state dynamics as

$$\begin{aligned}
\operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})^{-1} \Delta X_t^{***} &= \operatorname{diag}(\lambda^{\mathbb{Q}}) \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})^{-1} X_{t-1}^{***} + (A^{\mathbb{Q}})^{-1} \Sigma_X \epsilon_t^{\mathbb{Q}}, \\
\Delta X_t^{***} &= \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}}) \operatorname{diag}(\lambda^{\mathbb{Q}}) \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})^{-1} X_{t-1}^{***} \\
&\quad + \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})(A^{\mathbb{Q}})^{-1} \Sigma_X \epsilon_t^{\mathbb{Q}}, \\
\Delta X_t^{***} &= \operatorname{diag}(\lambda^{\mathbb{Q}}) X_{t-1}^{***} + \Sigma_{X^{***}} \epsilon_t^{\mathbb{Q}},
\end{aligned}$$

where

$$\Sigma_{X^{***}} = \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})(A^{\mathbb{Q}})^{-1} \Sigma_X$$

or

$$\Sigma_X = A^{\mathbb{Q}} \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})^{-1} \Sigma_{X^{***}} = V \Sigma_{X^{***}},$$

where $V = A^{\mathbb{Q}} \operatorname{diag}({\rho_{1X}}' A^{\mathbb{Q}})^{-1}$.

And with

$$X_t^{***} = \text{diag}({\rho_{1X}}' A^{\mathbb{Q}}) X_t^{**}$$

and

$$X_t^{**} = (A^{\mathbb{Q}})^{-1}[X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}]$$

the rotation between X_t and X_t^{***} is given by

$$\begin{aligned} X_t^{***} &= \text{diag}({\rho_{1X}}' A^{\mathbb{Q}})(A^{\mathbb{Q}})^{-1}[X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}] \\ &= V^{-1}[X_t + (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}]. \end{aligned}$$

Finally, the dividend growth equation can be rewritten as

$$\begin{aligned} \Delta d_t &= \delta_{0X} + {\delta_{1X}}' X_t \\ &= \delta_{0X} + {\delta_{1X}}'(V X_t^{***} - (K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}) \\ &= (\delta_{0X} - {\delta_{1X}}'(K_{1X}^{\mathbb{Q}})^{-1} K_{0X}^{\mathbb{Q}}) + {\delta_{1X}}' V X_t^{***}. \end{aligned}$$

Hence, overall, the canonical form of X_t is given by

$$\Delta X_t = \text{diag}(\lambda^{\mathbb{Q}}) X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}},$$

$$r_t = r_{\infty}^{\mathbb{Q}} + \iota' X_t,$$

$$\Delta d_t = \delta_{0X} + \delta_{1X}' X_t.$$

A1.3 Proof of Proposition 2

Given the canonical form obtained from Proposition 1, i.e.

$$\Delta X_t = \text{diag}(\lambda^{\mathbb{Q}})X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}},$$

$$r_t = r_\infty^{\mathbb{Q}} + \iota' X_t,$$

$$\Delta d_t = \delta_{0X} + \delta_{1X}' X_t,$$

and the rotational relationship between $\mathcal{P}_t$ and X_t, i.e.

$$\mathcal{P}_t = A + BX_t \quad (X_t = B^{-1}\mathcal{P}_t - B^{-1}A),$$

substitute $X_t = B^{-1}\mathcal{P}_t - B^{-1}A$ into the canonical form of X_t to obtain

$$\begin{aligned}
(B^{-1}\mathcal{P}_t - B^{-1}A) - (B^{-1}\mathcal{P}_{t-1} - B^{-1}A)& \\
= \text{diag}(\lambda^{\mathbb{Q}})(B^{-1}\mathcal{P}_{t-1} - B^{-1}A) + \Sigma_X \epsilon_t^{\mathbb{Q}}&, \\
r_t = r_\infty^{\mathbb{Q}} + \iota'(B^{-1}\mathcal{P}_t - B^{-1}A)&, \\
\Delta d_t = \delta_{0X} + \delta_{1X}'(B^{-1}\mathcal{P}_t - B^{-1}A)&,
\end{aligned}$$

which simplifies to

$$\Delta \mathcal{P}_t = -B\,\text{diag}(\lambda^{\mathbb{Q}})B^{-1}A + B\,\text{diag}(\lambda^{\mathbb{Q}})B^{-1}\mathcal{P}_{t-1} + B\Sigma_X \epsilon_t^{\mathbb{Q}},$$

$$r_t = r_\infty^{\mathbb{Q}} - \iota' B^{-1}A + \iota' B^{-1}\mathcal{P}_t,$$

$$\Delta d_t = \delta_{0X} - \delta_{1X}' B^{-1}A + \delta_{1X}' B^{-1}\mathcal{P}_t.$$

In order for the canonical form of $\mathcal{P}_t$ to be consistent with the canonical form of X_t, the following parameter restrictions follow:

$$
\begin{aligned}
K_{1\mathcal{P}}^{\mathbb{Q}} &= B\,\mathrm{diag}(\lambda^{\mathbb{Q}})B^{-1}, \\
K_{0\mathcal{P}}^{\mathbb{Q}} &= -K_{1\mathcal{P}}^{\mathbb{Q}}A, \\
\rho_{1\mathcal{P}} &= (B^{-1})'\iota, \\
\rho_{0\mathcal{P}} &= r_{\infty}^{\mathbb{Q}} - A'\rho_{1\mathcal{P}}, \\
\delta_{1\mathcal{P}} &= (B^{-1})'\delta_{1X}, \\
\delta_{0\mathcal{P}} &= \delta_{0X} - A'\delta_{1\mathcal{P}}.
\end{aligned}
$$

Appendix 2

A2.1 Impulse responses

To derive the impulse responses of the yields from shocks to the state factors $X_t = (\mathcal{P}_t', \mathcal{M}_t')'$, consider the physical dynamics of X_t in (3.1):

$$X_t = K_{0X}^{\mathbb{P}} + K_{1X}^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}},$$

which can be written as an implied Wold MA(∞) representation:

$$X_t = (I - K_{1X}^{\mathbb{P}})^{-1} K_{0X}^{\mathbb{P}} + \sum_{i=0}^{\infty} P_i \epsilon_{t-i}^{\mathbb{P}},$$

where $P_i = (K_{1X}^{\mathbb{P}})^i \Sigma_X$.

Moreover, given the bond yield equation (3.11),

$$y_{nt}^{\mathrm{b}} = A_n + B_n' \mathcal{P}_t,$$

the bond yield at maturity n, y_{nt}^{b}, can be written as

$$\begin{aligned} y_{nt}^{\mathrm{b}} &= A_n + \begin{bmatrix} B_n' & 0' \end{bmatrix} X_t \\ &= A_n + \begin{bmatrix} B_n' & 0' \end{bmatrix} (I - K_{1X}^{\mathbb{P}})^{-1} K_{0X}^{\mathbb{P}} + \sum_{i=0}^{\infty} \psi_i^n \epsilon_{t-i}^{\mathbb{P}}, \end{aligned}$$

where $\psi_i^n = [\, B_n' \; 0' \,] P_i$. That is, the yield on an n-period zero-coupon bond y_{nt}^{b} is a linear combination of current and lagged values of ϵ_t, where the row vectors ψ_i^n are a function of B_n.

The vector ψ_i^n is the impulse responses for the n-period yield at horizon i for shocks to the state vector X_t at time 0. For k yields of maturities $n_1, \ldots, n_k$, we can stack the coefficients of each yield to write

$$Y_t = A + \begin{bmatrix} B' & 0' \end{bmatrix} (I - K_{1X}^{\mathbb{P}})^{-1} K_{0X}^{\mathbb{P}} + \sum_{i=0}^{\infty} \Psi_i \epsilon_{t-i},$$

where $Y_t = (y_{n_1 t}^{\mathrm{b}} \; \ldots \; y_{n_k t}^{\mathrm{b}})'$ and the jth row of Ψ_i is $\psi_i^{n_j}$.

A2.2 Variance decompositions

Working with the MA(∞) representation of the yields, the error of the optimal h-step-ahead forecast at time t, $\hat{Y}_{t+h|t}$, is

$$\hat{Y}_{t+h|t} - Y_{t+h} = \sum_{i=0}^{h-1} \Psi_i \epsilon_{t+h-i}.$$

Let $\Psi_{jk,i}$ denote the element in row j, column k of Ψ_i. Then

$$\hat{Y}_{t+h|t}^j - Y_{t+h}^j = \sum_{k=1}^{K} (\Psi_{jk,0} \epsilon_{t+h}^k + \cdots + \Psi_{jk,h-1} \epsilon_{t+1}^k).$$

Denoting the mean squared error of $\hat{Y}_{t+h|t}^j$ as MSE($\hat{Y}_{t+h|t}^j$), we have

$$\mathrm{MSE}(\hat{Y}_{t+h|t}^j) = \sum_{k=1}^{K} (\Psi_{jk,0}^2 + \cdots + \Psi_{jk,h-1}^2).$$

The contribution $\Omega_{jk,h}$ of the kth factor to the MSE of the h-step ahead forecast of the jth yield is

$$\Omega_{jk,h} = \frac{\sum_{i=0}^{h-1} \Psi_{jk,i}^2}{\text{MSE}(\hat{Y}_{t+h|t}^j)},$$

which decomposes the forecast variance at horizon h of the jth yield to each of the K state factors.