EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1045 / APRIL 2009

THE TERM STRUCTURE OF EQUITY PREMIA IN AN AFFINE ARBITRAGE-FREE MODEL OF BOND AND STOCK MARKET DYNAMICS

by Wolfgang Lemke
and Thomas Werner

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1045 / APRIL 2009

THE TERM STRUCTURE OF EQUITY PREMIA IN AN AFFINE ARBITRAGE-FREE MODEL OF BOND AND STOCK MARKET DYNAMICS [1]

by Wolfgang Lemke

and Thomas Werner [2]

In 2009 all ECB publications feature a motif taken from the €200 banknote.

1 The views expressed in this paper do not necessarily represent the view of the European Central Bank or the Deutsche Bundesbank. We thank Gonzalo Camba-Méndez, Manfred Kremer, participants of the workshop 'Estimating the cost of equity and the equity premium: purposes, methods and implications for monetary policy analysis' (organized by the ECB and the Bank of England) and seminar participants at WHU Otto Beisheim School of Management for useful suggestions and discussion. Thanks to Renate Dreiskena for data assistance.

2 European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; e-mail: wolfgang.lemke@ecb.europa.eu (corresponding author, also Deutsche Bundesbank); thomas.werner@ecb.europa.eu

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Website
http://www.ecb.europa.eu

Fax
+49 69 1344 6000

ISSN 1725-2806 (online)

CONTENTS

Abstract

We estimate time-varying expected excess returns on the US stock market from 1983 to 2008 using a model that jointly captures the arbitrage-free dynamics of stock returns and nominal bond yields. The model nests the class of affine term structure (of interest rates) models. Stock returns and bond yields as well as risk premia are affine functions of the state variables: the dividend yield, two factors driving the one-period real interest rate and the rate of inflation. The model provides for each month the `term structure of equity premia', i.e. expected excess stock returns over various investment horizons. Model-implied equity premia decrease during the `dot-com' boom period, show an upward correction thereafter, and reach highest levels during the financial turmoil that started with the 2007 subprime crisis. Equity premia for longer-term investment horizons are less volatile than their short-term counterparts.

Keywords: Equity premium, affine term structure models, asset pricing

JEL Classification: E43, G12

Non-technical summary

The equity premium measures the additional return that an investor expects to gain by holding equity (or a portfolio of equities) instead of investing his funds into a risk-free bond. Since equity constitutes a major asset class, the equity premium is often considered a general yardstick for investors' overall desired risk compensation. Given the forward-looking nature of the equity premium, it is immediately clear that it is not a directly observable measure. A large academic literature has proposed a host of approaches to quantify excess equity returns. One of the simplest approaches is to consider average *realized* stock returns over several months or years and to compare these to the returns of bond investments over comparable horizons. While this approach comes with the advantage of being independent of a particular model, it does not – by construction – provide a timely estimate of the equity premium. In order to obtain estimates for the excess *expected* stock returns at a particular point in time, one either has to ask investors on their respective return expectations (survey approach) or use a model.

One of the most widely used frameworks in this respect is the so-called dividend-discount model. It is based on the basic idea that stock prices reflect discounted expected future dividends. The different versions of the dividend-discount model are characterized by the way in which they quantify expected future dividends. The easiest approach is to assume that future dividend growth rates are all constant ('Gordon growth model'). Alternatively, dividend growth rates over future time periods are extracted from survey information about future stock earnings. Having a path of expected future dividends at hand, and given the observed stock price and dividend at a given point in time, one then solves for the discount rate that equates the current price to the discounted stream of expected future dividends. Subtracting from this discount rate a long-term government bond yield provides an estimate of the equity premium at this point in time. Noteworthy, the employed stock return is assumed to be constant for all future horizons. In this sense, it has the flavour of a yield to maturity (or internal rate of return), as it is commonly used for measuring the returns of coupon-bearing bonds. However, by discounting all future dividends by this same rate of return, the term and risk structure of future expected stock returns is ignored. In fact, heuristically, future stock returns used for discounting dividends occurring at different future horizons may be understood as risk-adjusted interest rates for the respective horizons. Hence, one should be able to account for a term structure of expected stock returns, which is potentially non-flat, implying that expected annualized returns may differ over different investment horizons.

In this paper, we propose a model for the equity premium that takes the risk and term structure of required stock returns seriously. It does so by pricing bonds and stocks in a joint arbitrage-free framework, using a common stochastic discount factor (or 'pricing kernel') which prices both stocks and bonds of all maturities. Hence, the presented model captures the joint dynamics of bond and stock returns, where the no-arbitrage condition provides a set of restrictions on their comovements and on how they depend on economic driving factors. In our model, two (unobservable) factors drive the short-term real interest rate; a third factor is inflation, which is required to price nominal bonds; and the dividend yield – additionally required for the pricing of stocks – is included as a fourth factor. Nominal bond yields of all maturities are affine (linear plus intercept) functions of the real-rate factors and inflation. In fact, our model encompasses the well-established class of affine term structure (of interest rates) models, which have become a workhorse tool for estimating bond risk premia. This affine property is inherited for equity, implying that expected stock returns as well as equity premia for all horizons are affine functions of the factors as well.

Our model allows to derive at each point in time the whole 'term structure of equity risk premia'. That is, for an arbitrary investment horizon, it provides the excess expected return of equity over the model-implied real interest rate with the corresponding time to maturity.

Regarding the empirical contributions, we estimate the model on US data from 1983 to 2008. The model gives a good fit to government bond yields, and it implies bond yield risk premia, which are comparable in size and dynamics to those obtained in the affine term structure (of interest rates) literature. Hence, enhancing an otherwise standard affine term structure model to price also common stock does not adversely affect its capability of capturing salient bond market features.

As a cross-check, the model-implied equity premia are compared to the equity premium implied by the three-stage dividend discount model, which is a well-established benchmark, especially among practitioners. It turns out that the model-implied premia, exhibit a marked comovement with the equity premium obtained from the dividend-discount model, which supports the empirical validity of our model. At the same time, the equity premia implied by the model can be interpreted as measures of required risk compensation for holding equity over well-defined investment horizons, whereas the equity premium from the dividend-discount model rather represents an average (over a set of horizons) excess return.

After hovering around a level somewhat above three percent, our model-implied equity premia show a distinct downward trend starting in 1995. This is recorded across the whole range of investment horizons. The lowest levels of expected excess returns are reached by the year 2000 when the 'dot-com' euphoria reached its climax. The following normalization of equity risk premia eventually reaches a new turning point at the beginning of 2003, after which equity risk premia decline again, but especially for shorter investment horizons. Finally, with the onset of the financial turmoil started by the subprime crisis in summer 2007, equity risk premia of all horizons climb up to the highest levels recorded over our sample period.

Overall, equity premia of all investment horizons show a strong comovement but those for longer horizons turn out to be less volatile in our sample. For instance, the annualized one-year expected excess returns takes values in the range of minus 0.1 percent (at the height of the dot-com episode) to 6.2 percent (at end-2008), while the model-implied hundred-year equity premium ranges between 1.5 and 4 percent.

1 Introduction

Absence of arbitrage is the key condition underlying models of asset pricing. This condition is known to be satisfied if there is a common 'pricing kernel' for valuating the payoffs of all assets.[1] As argued in Cochrane (2001), specifying the joint evolution of the pricing kernel and asset payoffs constitutes a unifying framework for modeling the price and return profiles of any family of assets. Following this approach, our paper proposes a discrete-time arbitrage-free model that captures the joint dynamics of two major asset classes: dividend-paying equity and government bonds of different maturities. The pricing kernel depends on four factors: the rate of inflation, the dividend yield and two additional factors driving the real rate of interest. The main application of the model is the derivation of time-varying bond and equity risk premia within a common framework. In particular, equity premia, i.e. expected excess stock returns, can be extracted for arbitrary investment horizons. The model is estimated on monthly US data from 1983 to 2008.

As a convenient property of our model, arbitrage-free bond yields and stock returns as well as bond and equity risk premia all result as affine functions of the state variables. In fact, the model nests the popular class of essentially affine term structure (of interest rates) models, on which there is a large and still growing literature.[2] The research in this area was rather successful in showing that time-varying bond market risk premia can explain observed shortcomings of the expectations hypothesis of the yield curve. As equity valuation essentially requires discounting future dividend cash flows, the integration of stock and bond pricing in a single affine framework could be seen as the natural step forward.

Somewhat surprisingly, however, there is relatively little literature on integrating these two asset classes in a common arbitrage-free framework. One of the first papers in this respect is Bekaert and Grenadier (2001), who use the the same modeling approach as for affine term structure (of interest rates) models, but include the dividend growth associated with a representative portfolio of equities into the state vector. Bond prices are still exponentially affine functions of the state variables but dividend-scaled stock prices are infinite sums of such functions. Hence, while bond yields have an affine representation as usual, stock returns have not. The same holds for the model by Lettau and Wachter (2007).

As an alternative, Mamaysky (2002) proposed a continuous-time affine model which includes the dividend yield (dividend over ex-dividend stock price) as state variable instead of the dividend growth. This comes with the advantage that stock prices have an exponentially affine closed-form solution. Equity returns are consequently affine functions

[1] See Harrison and Kreps (1979) and Cochrane (2001).

[2] See Duffie and Kan (1996) and Dai and Singleton (2000). Chapter 13 of Singleton (2006) provides an overview and contains several references.

of the state variables in this framework. Besides the fact that exact affine pricing equations for bond yields and equity returns greatly facilitate the estimation of the model, the approach chosen by Mamaysky has the advantage that it does not rely on the forecastability of dividend growth. In fact, as argued by Cochrane (2008) recently, the empirical evidence does not support the forecastability of this variable for US data. At the same time, dividend yields, albeit rather persistent, show evidence of being mean-reverting.

The paper contributes to the literature on affine asset pricing models and to the empirical finance literature on the US stock market.

First, from a technical point of view, the paper provides a general discrete-time valuation framework with closed-form affine solutions for bond yields and stock returns. It thereby encompasses the class of multi-factor affine term structure (of interest rate) models. The papers by d'Addona and Kind (2006) and Li (2002) also develop discrete time models with closed-form affine solutions for stock returns but assume more restricted factor dynamics and constant rather than time-varying market prices of risk. This implies that all bond and equity premia are constant over time. In our paper, in contrast, the market price of risk specification is very general using the essentially affine approach proposed by Duffee (2002). This allows for various risk factors to affect the equity risk premium. Our model allows to derive at each point in time the whole 'term structure of equity risk premia'. That is, for an arbitrary investment horizon, it provides the excess expected return of equity over the model-implied real interest rate with the corresponding time to maturity.

Second, regarding the empirical analysis of bond and stock markets using an affine framework, our paper is novel in focussing on the time series of the equity risk premium. The aforementioned analyses by d'Addona and Kind (2006) and Li (2002), in contrast, gear to understanding the correlation of bond and stock returns.

Concerning the empirical results, our model gives a good fit to US government bond yields, and it implies bond yield risk premia, which are comparable in size and dynamics to those obtained in the affine term structure (of interest rates) literature. Hence, enhancing an otherwise standard affine term structure model to price also common stock does not adversely affect its capability of capturing salient bond market features.

As a cross-check, the model-implied equity premia are compared to the equity premium implied by the three-stage dividend discount model, which is a well-established benchmark, especially among practitioners. It turns out that the model-implied premia, exhibit a marked comovement with the equity premium obtained from the dividend-discount model, which supports the empirical validity of our model. At the same time, the equity premia implied by our model can be interpreted as measures of required risk compensation for holding equity over well-defined investment horizons, whereas the equity premium from the dividend-discount model rather represents an average (over a set of horizons) excess return.

After hovering around a level somewhat above three percent, our model-implied equity premia show a distinct downward trend starting in 1995. This is recorded across the whole range of investment horizons. The lowest levels of expected excess returns are reached by the year 2000 when the 'dot-com' euphoria reached its climax. The following normalization of equity risk premia eventually reaches a new turning point at the beginning of 2003, after which equity risk premia decline again, but especially for shorter investment horizons. Finally, with the onset of the financial turmoil started by the subprime crisis in summer 2007, equity risk premia of all horizons climb up to the highest levels recorded over our sample period.

Overall, equity premia of all investment horizons show a strong comovement but those for longer horizons turn out to be less volatile in our sample. For instance, the annualized one-year expected excess returns takes values in the range of minus 0.1 percent (at the height of the dot-com episode) to 6.2 percent (at end-2008), while the model-implied hundred-year equity premium ranges between 1.5 and 4 percent.

The paper is structured as follows: the next section develops the joint stock-bond-pricing model, where lengthier derivations are delegated to the appendix. Section 3 compares the arbitrage-free model of this paper to the popular dividend discount model. Section 4 explains how the model is cast into state space form, documents the parameter restrictions used for estimation and presents the data. Section 5 contains the empirical results: first, parameter estimates and the empirical fit are reported. This is followed by a discussion of the estimated series of bond and equity risk premia. It also includes an interpretation of the 'term structure of equity risk premia'. Section 6 concludes and provides perspectives for future research.

2 An affine arbitrage-free model of bond and stock market dynamics

We specify a model for the joint arbitrage-free dynamics of bond yields and stock returns. Time is discrete and the unit time interval can be understood as one month. We will derive the pricing equations for nominal bonds of arbitrary maturity and one dividend-paying stock. This is motivated by the fact that we will use several nominal zero-coupon yields and one broad-based stock index for estimating the model. Hence, while the generalization to a family of dividend-paying securities would be straightforward, we will focus in the following on one stock that will be interchangeably be referred to as 'the stock' or 'the stock index'.

The core component of the model is a pricing kernel that prices assets that pay off in real terms. By also specifying the dynamics of inflation, we obtain the pricing kernel for nominal assets, which is required to compute the arbitrage-free dynamics of the term structure of nominal bond yields. Besides inflation, there are three other risk factors in the model, two of them driving the one-month risk-free real rate and a factor representing the payout yield of the stock index. The term 'payout yield' refers to the payout of the stock index divided by its price. Viewed narrowly, this is tantamount to the dividend yield. However, as listed companies also have other measures at their disposal to let stock holders participate in profits (e.g. stock buy backs), the 'payout yield' subsumes all payments to investors, of which dividends may only be a part. Nevertheless, for simplicity, we will interchangeably use the term 'dividend yield' for the same variable, and likewise use the word 'dividends' for what rather refers to the total payout to equity holders.

The solution of the model results as a system comprising the linear dynamics of the factor process as well as a a set of affine equations relating bond yields and stock returns to the factors. Hence, the resulting system encompasses the class of affine term structure (of interest rates) models.

In the following, we first describe the dynamics of the factor process, which is followed by a specification of the pricing kernels. We then turn to the derivation of the arbitrage-free term structure dynamics of nominal bonds and the pricing problem for equity. After that, bond and equity risk premia are derived.

2.1 Factor process

Let $X_t := (\pi_t, \gamma_t, L_{1t}, L_{2t})'$ denote a vector that contains inflation π_t, a payout-yield factor γ_t as well as two additional factors L_{1t} and L_{2t} that constitute the one-period real interest rate r_t. More precisely, π_t is the logarithmic month-on-month change of the level Π_t of a consumer price index, i.e. $\pi_t := \ln \Pi_t - \ln \Pi_{t-1}$. The payout-yield factor is defined as $\Gamma_t := (1 + \frac{D_t}{V_t})$, where D_t is the dividend of the stock in one-period terms and V_t is the ex-dividend stock price. The factor vector contains the log of that, $\gamma_t := \ln \Gamma_t$. Thus, γ_t approximates the dividend yield $\frac{D_t}{V_t}$. Finally, the real interest rate is an affine function of L_{1t} and L_{2t},

$$r_t = \delta_0 + \delta_1' X_t, \quad \delta_1 = (0, 0, \delta_{1,3}, \delta_{1,4})'. \tag{2.1}$$

As in the affine term structure (of interest rates) literature, the factor dynamics is specified as a stationary VAR(1),

$$X_{t+1} = a + \mathcal{K} X_t + \Sigma \, \eta_{t+1}, \quad \eta_t \underset{i.i.d.}{\sim} N(0, I) \tag{2.2}$$

where a, $\mathcal{K}$ and Σ are a vector and matrices, respectively, of appropriate dimensions. The empirical counterparts of the elements of the factor vector will be discussed in section 4 below.

2.2 Real pricing kernel

We define the real pricing kernel, or stochastic discount factor (SDF), M_t as

$$M_{t+1} = \exp(-r_t)\frac{\xi_{t+1}}{\xi_t} \tag{2.3}$$

where r_t is the real one-month interest rate, and the risk-adjustment term satisfies

$$\xi_{t+1} = \xi_t \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\eta_{t+1}\right), \tag{2.4}$$

with

$$\lambda_t = \lambda_0 + \Lambda_1 X_t \tag{2.5}$$

being the vector of market prices of risk, where λ_0 and Λ_1 are a vector and a matrix of parameters. That is, relations (2.3) to (2.5) allow all risk factors (real rate, inflation, dividend yield) to be priced, and the risk prices themselves are spanned by the factors. Hence, we take the same flexible approach as in most of the affine term structure (of interest rate) literature.[3]

For the log real stochastic discount factor, $m_t := \ln M_t$, we have

$$m_{t+1} = -\frac{1}{2}\lambda_t'\lambda_t - \delta_0 - \delta_1' X_t - \lambda_t'\eta_{t+1}. \tag{2.6}$$

Under the condition of no arbitrage, the price at time t of an asset i with real payoff Z_{t+1}^i in period $t+1$ satisfies

$$P_t^i = E_t\{M_{t+1}Z_{t+1}^i\}. \tag{2.7}$$

2.3 Nominal pricing kernel

Similarly, assets that pay off in nominal terms (i.e. in units of currency) are priced by a nominal SDF $\tilde{M}_t$, so their prices are given by

$$\tilde{P}_t^i = E_t\{\tilde{M}_{t+1}\tilde{Z}_{t+1}^i\}. \tag{2.8}$$

In the following, if not stated otherwise, a tilde on top of a variable (price, return, stochastic discount factor) will denote that it relates to nominal as opposed to real assets. The log nominal and the log real SDF are related by[4]

$$\tilde{m}_{t+1} = m_{t+1} - \pi_{t+1}. \tag{2.9}$$

[3] An alternative route is chosen by Lettau and Wachter (2007) who specify a separate process for the market price of dividend(-growth) risk. The dynamics of this variable is a simple AR(1), but its innovations are allowed to be correlated with the innovations of dividend growth and inflation.

[4] See, e.g., Campbell, Lo, and MacKinlay (1997).

Let δ_π denote a selection vector that picks inflation from the factor vector, i.e. $\pi_t = \delta'_\pi X_t$. Using (2.9) one obtains for the nominal log SDF $\tilde{m}_{t+1}$:

$$\begin{aligned}\tilde{m}_{t+1} &= m_{t+1} - \delta'_\pi X_{t+1} \\ &= m_{t+1} - \delta'_\pi(a + \mathcal{K}X_t + \Sigma\eta_{t+1}) \\ &= -\frac{1}{2}\lambda'_t\lambda_t - (\delta_0 + \delta'_\pi a) - (\delta'_1 + \delta'_\pi\mathcal{K})X_t - \lambda'_t\eta_{t+1} - \delta'_\pi\Sigma\eta_{t+1} \qquad (2.10)\end{aligned}$$

To see that the nominal and the real SDF have a perfectly analogous functional form, we define $\tilde{\lambda}_t := \lambda_t + \Sigma'\delta_\pi$. It satisfies

$$\tilde{\lambda}'_t\tilde{\lambda}_t = (\lambda_t + \Sigma'\delta_\pi)'(\lambda_t + \Sigma'\delta_\pi) = \lambda'_t\lambda_t + 2\delta'_\pi\Sigma\lambda_0 + 2\delta'_\pi\Sigma\Lambda_1 X_t + \delta'_\pi\Sigma\Sigma'\delta_\pi.$$

Replacing λ_t in (2.10) we obtain an expression for the log nominal SDF $\tilde{m}_{t+1}$ which is analogous in structure to (2.6),

$$\tilde{m}_{t+1} = -\frac{1}{2}\tilde{\lambda}'_t\tilde{\lambda}_t - \tilde{\delta}_0 - \tilde{\delta}'_1 X_t - \tilde{\lambda}'_t\eta_{t+1}. \qquad (2.11)$$

The mapping from the 'real' parameters to the 'nominal' parameters (with tilde) is given by:

$$\begin{aligned}\tilde{\lambda}_t &:= \lambda_t + \Sigma'\delta_\pi, \quad \text{thus} \quad \tilde{\lambda}_0 \equiv \lambda_0 + \Sigma'\delta_\pi, \quad \tilde{\Lambda}_1 \equiv \Lambda_1 \qquad (2.12)\\ \tilde{\delta}_0 &:= \delta_0 + \delta'_\pi(a - \Sigma\lambda_0) - \frac{1}{2}\delta'_\pi\Sigma\Sigma'\delta_\pi, \qquad (2.13)\\ \tilde{\delta}'_1 &:= \delta'_1 + \delta'_\pi(\mathcal{K} - \Sigma\Lambda_1). \qquad (2.14)\end{aligned}$$

In analogy to (2.6), where $r_t := \delta_0 + \delta'_1 X_t$ represents the real interest rate, $i_t := \tilde{\delta}_0 + \tilde{\delta}'_1 X_t$ in (2.11) represents the one-period nominal interest rate. One observes that the two are related as

$$r_t = i_t - \delta'_\pi(a + \mathcal{K}X_t) + \delta'_\pi\Sigma\lambda_t + \frac{1}{2}\delta'_\pi\Sigma\Sigma'\delta_\pi, \qquad (2.15)$$

hence, the real short rate equals its nominal counterpart minus expected inflation, plus a risk-premium (which is zero if λ_0 and Λ_1 are both zero) and a small Jensen inequality term.[5]

2.4 Pricing nominal zero-coupon bonds

Given the factor process and the real as well as the nominal pricing kernel, we can price real and nominal assets. For nominal zero-coupon bonds, relation (2.8) becomes

$$\tilde{P}^n_t = E_t\{\tilde{M}_{t+1}\tilde{P}^{n-1}_{t+1}\}, \qquad (2.16)$$

where $\tilde{P}^n_t$ is the price at time t of a zero-coupon bond maturing at time $t+n$, when it pays one unit of currency, i.e $\tilde{P}^0_{t+n} = 1$. As is well known, the chosen specifications

[5] For our estimation of the model it amounts to much less then 0.01%.

of the factor process, the SDF and the market price of risk imply that bond prices are exponentially-affine functions of the factors[6],

$$\tilde{P}_t^n = \exp\left[\bar{\tilde{A}}_n + \bar{\tilde{B}}_n' X_t\right], \tag{2.17}$$

where the coefficients $\bar{\tilde{A}}_n$ and $\bar{\tilde{B}}_n$ satisfy the following system of difference equations in n,

$$\bar{\tilde{A}}_n = \bar{\tilde{A}}_{n-1} + \bar{\tilde{B}}_{n-1}'(a - \Sigma\tilde{\lambda}_0) + \frac{1}{2}\bar{\tilde{B}}_{n-1}'\Sigma\Sigma'\bar{\tilde{B}}_{n-1} - \tilde{\delta}_0 \tag{2.18}$$

$$\bar{\tilde{B}}_n' = \bar{\tilde{B}}_{n-1}'(\mathcal{K} - \Sigma\tilde{\Lambda}_1) - \tilde{\delta}_1' \tag{2.19}$$

with initial conditions $\bar{\tilde{A}}_0 = 0$ and $\bar{\tilde{B}}_0 = 0_N$.[7] Hence, continuously-compounded bond yields, defined as $\tilde{y}_t^n := -\frac{\ln \tilde{P}_t^n}{n}$, will be affine functions of the factors

$$\tilde{y}_t^n = \tilde{A}_n + \tilde{B}_n' X_t, \tag{2.20}$$

where $\tilde{A}_n = -\frac{\bar{\tilde{A}}_n}{n}$ and $\tilde{B}_n = -\frac{\bar{\tilde{B}}_n}{n}$.

The pricing equation for real bond yields is completely analogous, thus

$$y_t^n = A_n + B_n' X_t, \tag{2.21}$$

where A_n and B_n satisfy (2.18) and (2.19) with all symbols carrying a tilde being replaced by the respective symbol without one.

2.5 Pricing dividend-paying stocks

Denote by D_t the real dividend of a stock paid at time t and by V_t the stock's real (ex-dividend) price at time t. Buying one unit of the stock at time t at a price of V_t entitles the stock holder to next period's dividend D_{t+1}, and the stock can then be sold for the next period's price V_{t+1}. Hence, the total payoff is given by $D_{t+1} + V_{t+1}$. Therefore, using (2.7), the stock price satisfies

$$V_t = E_t\left\{M_{t+1}(D_{t+1} + V_{t+1})\right\}. \tag{2.22}$$

Using the definition of the payout-yield factor $\Gamma_t := (1 + \frac{D_t}{V_t})$, this can be rewritten as

$$V_t = E_t\left\{M_{t+1}\Gamma_{t+1}V_{t+1}\right\}. \tag{2.23}$$

As derived in appendix A.1, this expectational difference equation has the solution[8]

[6] See, e.g., Ang and Piazzesi (2003).

[7] N denotes the dimension of the factor vector, i.e. here $N = 4$.

[8] Similar to the procedure used by Mamaysky (2002) for his continuous-time model, we start with the guess of an exponentially affine solution for the stock price and solve for the undetermined coefficients. An alternative approach is employed by d'Addona and Kind (2006) and Li (2002) who start with analytical solutions of pricing equations for n-period dividend paying securities and price equities by taking limits to infinity. It can be shown that the two approaches yield the same result: casting Li's model into our general framework generates the same expression for stock returns as in his paper. The proof is available on request.

$$V_t^* = \exp[c \cdot (t - t_0) + D' X_t], \tag{2.24}$$

where

$$\begin{aligned} D' &= \left[\delta_\gamma'(\mathcal{K} - \Sigma\Lambda_1) - \delta_1'\right] \cdot \left[I_N - (\mathcal{K} - \Sigma\Lambda_1)\right]^{-1}, & (2.25) \\ c &= \delta_0 - (\delta_\gamma + D)' a - \frac{1}{2}(\delta_\gamma + D)'\Sigma\Sigma'(\delta_\gamma + D) + (\delta_\gamma + D)'\Sigma\lambda_0, & (2.26) \end{aligned}$$

and t_0 is a free parameter. The vector δ_γ has the second element equal to one and zeros elsewhere, i.e. it picks the dividend yield from the state vector, $\gamma_t = \delta_\gamma' X_t$.

The arbitrage-free stock price consists of a deterministic exponential trend and a stochastic fluctuation around this trend. Note that the absolute magnitude of the stock price is not pinned down by the model. Hence, it is useful to think of the solution $V_t = V_t^*$ as describing the dynamics of an index in arbitrary units of measurements that can be altered via t_0.[9]

Gross one-period stock returns are given by

$$R_{t+1}^{(1)} = \frac{V_{t+1} + D_{t+1}}{V_t} = \frac{V_{t+1}(1 + \frac{D_{t+1}}{V_{t+1}})}{V_t}.$$

Accordingly, one-period log-returns equal the capital gain (change of ex-dividend log stock price), $\Delta v_{t+1} = c + D'\Delta X_{t+1}$, plus the next period's dividend yield:

$$r_{t+1}^{(1)} = \Delta v_{t+1} + \gamma_{t+1} = c + D'\Delta X_{t+1} + \delta_\gamma' X_{t+1}, \tag{2.27}$$

Thus, conditionally expected returns are affine functions of the state vector

$$\begin{aligned} E_t\, r_{t+1}^{(1)} &= c + (D' + \delta_\gamma')a + (D'(\mathcal{K} - I) + \delta_\gamma'\mathcal{K})X_t \\ &=: f_1 + F_1' X_t & (2.28) \end{aligned}$$

with obvious definitions of f_1 and F_1.[10] From (2.27), it follows immediately that the unconditionally expected stock return equals

$$E\, r_t^{(1)} = c + \mu_\gamma \tag{2.29}$$

where $\mu_\gamma := E\gamma_t$.

[9] Note that our solution is 'bubble-free' by assumption. In fact, (2.24) is not the only solution to the pricing relation (2.23): there is a wider family of solutions, for which the stock price does not only depend on the four factors but also on an unrelated random walk process. Such a solution would be characterized as a 'rational bubble'. While it would in principle be interesting to allow for the presence of bubbles for explaining stock returns, we decided to exclude them in this paper and to take a 'purely fundamental' view on stock pricing.

[10] For the derivation, it has been used that $E_t\, X_{t+1} = a + \mathcal{K}X_t$ and $E_t\, \Delta X_{t+1} = a + (\mathcal{K} - I)X_t$.

With $\lambda_0 = 0$ and $\Lambda_1 = 0$ (risk neutrality), we have[11]

$$r_{t+1}^{(1)} = r_t + (D' + \delta_\gamma')\Sigma\eta_{t+1} - J \tag{2.30}$$

where the variance (Jensen) term is $J = \frac{1}{2}(\delta_\gamma + D)'\Sigma\Sigma'(\delta_\gamma + D)$. Taking conditional expectations yields

$$E_t r_{t+1}^{(1)} = r_t - J. \tag{2.31}$$

That is, under risk neutrality the expected real stock return equals the real interest rate (plus a Jensen adjustment).

For defining multi-period returns, one has to make an assumption on how investors treat dividends that they receive during the considered investment horizon. One possibility is to assume that dividends are always reinvested in the stock that they are associated with.[12] That is, for an n-period horizon, the investor would buy the stock (say 100 units) at some time t for the ex-dividend price of V_t per share. He would then receive dividends in $t+1$, which he would use for buying new pieces of the stock for the then prevailing ex-dividend price V_{t+1} and so forth. His total payoff in the last period $(t+n)$ consists of the number of stocks carried over from period $t+n-1$ multiplied by $(D_{t+n} + V_{t+n})$, i.e. the dividend per share plus the ex-dividend stock price prevailing in the last period. This investment strategy is formally analyzed in appendix A.3. One obtains for n-period returns (scaled, i.e. in per-one-period terms)

$$r_{t+n}^{(n)} = \frac{1}{n}\left(v_{t+n} - v_t + \sum_{i=1}^{n}\gamma_{t+i}\right). \tag{2.32}$$

Thus, in analogy to the one-period case, they equal the n-period capital gain plus the average of dividend yields over the horizon considered. Since conditional expectations of v_t, v_{t+n} and the γ_{t+i} are all affine in X_t, conditional expectations of n-period returns have likewise an affine representation of the form:[13]

$$E_t r_{t+n}^{(n)} = f_n + F_n' X_t, \tag{2.33}$$

where

$$\begin{aligned}
f_n &= c + \frac{1}{n}D'R(\mathcal{K},n)\mathcal{K}^{-1}a + \frac{1}{n}\delta_\gamma'(I-\mathcal{K})^{-1}[n\cdot I - R(\mathcal{K},n)]a,\\
F_n' &= \frac{1}{n}D'(I-\mathcal{K}^{-1})R(\mathcal{K},n) + \frac{1}{n}\delta_\gamma' R(\mathcal{K},n),\\
\text{with } R(\mathcal{K},n) &= (I-\mathcal{K})^{-1}\mathcal{K}(I-\mathcal{K}^n).
\end{aligned}$$

Note that this implies at each time t a 'term structure of expected stock returns'. Finally, *unconditionally* expected n-period stock returns are independent of n and equal

[11] As shown in Appendix A.2.

[12] Taking again the perspective that the stock considered here can be conveniently considered as an index, the assumption implies that all receipts are reinvested into the index.

[13] See appendix A.4

to (2.29), which directly follows from taking unconditional expectations of (2.32) and exploiting stationarity of the factor process X_t. That is, the term structure of unconditional expectations of stock returns is flat.[14]

2.6 Risk premia

The model implies the dynamics of equity and bond yield risk premia. We define the one-period equity risk premium *ERP* at time t as the expected excess one-period log return - as defined in (2.27) - over the one-period real bond yield,

$$ERP_t^{(1)} := E_t r_{t+1}^{(1)} - r_t. \tag{2.34}$$

As both the one-period real interest rate and the expected stock return are affine functions of the state vector, the equity risk premium inherits this convenient property:

$$\begin{aligned} ERP_t^{(1)} &= f_1 + F_1' X_t - \delta_0 - \delta_1' X_t \\ &=: g_1 + G_1' X_t. \end{aligned}$$

The n-period *ERP* can be defined as the difference of expected n-period stock returns, as defined above, and the n-period real bond yield:

$$\begin{aligned} ERP_t^{(n)} &= E_t r_{t+n}^{(n)} - y_t^n \\ &= f_n + F_n' X_t - A_n - B_n' X_t \\ &=: g_n + G_n' X_t, \end{aligned} \tag{2.35}$$

again an affine function of the state vector.

For a given point in time t, (2.35) defines a 'term structure of equity risk premia'. Given the model parameters, it can take on a variety of shapes (upward or downward sloping, hump-shaped) depending on the realization of the state vector X_t.

From the fact that unconditional expectations of stock returns are independent of n, the shape of the unconditional expectation of the term structure of *ERP*s depends solely on the shape of the term structure of unconditional expectations of real bond yields. If this is upward-sloping, the term structure of unconditional means of ERPs is downward sloping, since the term structure of unconditional expectations of stock returns is flat.

As stated above, our model nests the popular class of affine term structure (of interest rates) models. One of their most important applications is the provision of nominal term premia, i.e. the differences between nominal long-term bond yields and their hypothetical counterparts that would prevail under the expectations hypothesis. Thus, it is a useful

[14] At first sight, the representation (2.33) of conditional return expectations may suggest that unconditional expectations of stock returns depend on n, i.e. the term structure of unconditionally expected returns is not flat. However, it can be shown that when X_t in (2.33) is replaced by its unconditional expectation, the n-dependent expressions cancel.

validation exercise for our encompassing model to compare the resulting term premia to those stemming from more specialized 'bond-yield-curve-only' models. The n-period nominal term premium (or yield risk premium YRP) is defined as:

$$YRP_t^n = \tilde{y}_t^n - \frac{1}{n} E_t \left\{ \sum_{i=0}^{n-1} i_{t+1} \right\}. \tag{2.36}$$

Again, since our model implies that arbitrage-free bond yields as well as current and expected nominal short rates are affine functions of the state vector, yield risk premia (and likewise forward premia and excess expected holding-period returns) are also affine functions of X_t.[15]

3 The three-stage dividend discount model as a benchmark for comparison

In the empirical application, the equity premium from the widely-used dividend discount model will be employed as a yardstick for comparison with our estimated equity risk premium. Therefore, the following provides a summary description of the dividend-discount model as well as a short characterization of similarities and differences between that model and the model introduced in this paper.

3.1 A short description of the dividend discount model

In the dividend-discount model, the stock price is the sum of discounted expected future dividends

$$V_t = \sum_{i=1}^{\infty} \left(\frac{1}{1 + \overline{roe}_t} \right)^i E_t D_{t+i} \tag{3.1}$$

where $\overline{roe}_t$ is the one-period required rate of return, which is taken as constant for all future periods from time t henceforth.

For extracting a risk premium measure for stocks, the dividend-discount model takes the observed stock price as given, uses some assumptions concerning future dividend growth and solves for the discount rate $\overline{roe}$. In the last step, one then subtracts from $\overline{roe}$ a risk-free rate (usually a long-term government bond yield, say y_t^m) and treats the difference as the equity risk premium:

$$\overline{ERP}_t^{DDM} = \overline{roe}_t - y_t^m \tag{3.2}$$

Hence, given the observed stock price (index) V_t, the only ingredient needed to back out $\overline{roe}_t$ from (3.1) is the sequence of expected future dividends. Equivalently, the equation

[15] See, e.g., Hördahl, Tristani, and Vestin (2006) for the various definitions of bond-related risk premia

can be written in terms of the current dividend D_t and future dividend *growth* rates,

$$V_t = D_t \sum_{i=1}^{\infty} \left(\frac{1}{1+\overline{roe}_t} \right)^i E_t \left\{ \prod_{j=1}^{i} (1+g_{t+j}) \right\}, \quad 1+g_{t+j} = \frac{D_{t+j}}{D_{t+j-1}}, \tag{3.3}$$

hence inferring the equity premium requires an assumption on expected future dividend growth rates. The simplest approach is to assume these growth rates to be constant for all future horizons from t onwards, $g_{t+j} = \bar{g}_t$, which is the famous Gordon growth model endowed with some quantification of $\bar{g}_t$.

A popular refinement used in practice is the so-called three-stage dividend discount model.[16] The version employed by many practitioners and central banks uses IBES (Institutional Brokers Estimate System) forecasts of *earnings* growth rates as a central input. These survey figures are understood as 'long-run' forecasts relating to a time horizon of 'three to five years', which for the purpose of estimating the equity premium is usually taken to correspond to a four-year horizon. Furthermore, it is assumed that the ratio of dividends to earnings is roughly constant, so that earnings growth forecasts proxy well for dividend growth forecasts. Denote this expected growth rate to be plugged into (3.3) over the first four years as g_t^{IBES}. For the very long run, say from twelve years henceforth, a constant dividend growth rate g^{LR} is used. This is often equated with some (ad hoc) long-run dividend growth assumption. For the time period of eight years ('second stage') between the four years , for which the IBES forecast is used ('first stage'), and the time starting after twelve years ('third stage'), the dividend growth rate is linearly interpolated between g_t^{IBES} and g^{LR}. Under these assumptions, the stock-valuation formula (3.3) becomes

$$V_t = \frac{D_t}{\overline{roe}_t - g^{LR}} \left((1+g^{LR}) + 8 \cdot (g_t^{IBES} - g^{LR}) \right), \tag{3.4}$$

Note that (3.4) becomes the formula for the Gordon model if $g_t^{IBES} = g^{LR}$, i.e. when the expected dividend growth rate is assumed constant in all three stages.

Using equation (3.4), one can explicitly solve for the return on equity $\overline{roe}_t$ implied by the three-stage dividend discount model:

$$\overline{roe}_t = \frac{D_t}{V_t} \cdot \left((1+g^{LR}) + 8 \cdot (g_t^{IBES} - g^{LR}) \right) + g^{LR}. \tag{3.5}$$

For the empirical study below, the return on equity $\overline{roe}_t$ in formula (3.5) is computed using the following inputs: for $\frac{D_t}{V_t}$ the Datastream dividend yield for the US S&P 500 stock index is used. For the medium-term dividend growth rate expectations g_t^{IBES}, the IBES forecast for a horizon of three to five years ahead is employed, from which Consensus Forecast inflation expectations for a comparable horizon are subtracted to convert it to real terms. The constant long-run dividend growth rate g^{LR} is set to 3.5%. To arrive

[16] See Fuller and Hsia (1984) and the exposition of the simplified version by Panigirtzoglou and Scammell (2002).

at the equity premium the ten-year constant-maturity real yield provided by the Federal Reserve is subtracted from $\overline{roe}_t$.

3.2 A comparison of the arbitrage-free model and the dividend-discount model

The two models share the property that stock prices are represented as discounted sums of future dividends. However, the models differ in their respective notion of a 'discount factor'. In the arbitrage-free model, the discount factor is inside the expectations operator, implying that the stock price is the sum of *expected discounted* future dividends

$$V_t = \sum_{i=1}^{\infty} E_t \left\{ \bar{M}_{t+i} D_{t+i} \right\} \tag{3.6}$$

where $\bar{M}_{t+i} = M_{t+1} \cdot \ldots \cdot M_{t+i}$.[17]

The stock price in the dividend-discount model, in contrast, is given by the sum of *discounted expected* future dividends, see (3.1). Hence, it neglects the term structure and intertemporal risk structure of discount rates. Accordingly, the derived return on equity $\overline{roe}_t$ is arguably similar in nature to a 'yield to maturity'.[18] Thus, the resulting equity risk premium rather represents an *average* over the whole set of future horizons, while our approach yields expected excess returns for well-defined horizons.

Another difference is the fact that the (three-stage) dividend discount model uses survey information on dividend growth rates at least for the 'first stage', while our model does not. Rather, all conditional expectations on future dividend yields, stock prices and interest rates are implied by the arbitrage-free model dynamics; accordingly, they are fully determined by current state variables.

4 Estimation approach and data

4.1 The empirical model in state space form

Regarding the estimation of our model, the advantage of choosing dividend yield as opposed to dividend growth as part of the state vector becomes evident. Unlike with the approach of, e.g., Bekaert and Grenadier (2001), stock returns in our model are affine function of the state vector. Hence, as common for affine term structure (of interest rates) models, the combined stock-bond model can be estimated in a state space framework.[19]

---

[17] Note that equation (3.6) results as the forward solution of (2.22), and utilizing the transversality condition $\lim_{n\to\infty} E_t \left[\left(\prod_{i=1}^{n} M_{t+i} \right) V_{t+n} \right] = 0$.

[18] As one would derive from the price of a coupon-bearing bond and information about its future coupon and principal payments.

[19] See, e.g., Lemke (2006) for an overview. The model by Bekaert and Grenadier (2001) is estimated using an iterated GMM method. d'Addona and Kind (2006) estimate the factor process parameters by

The system of measurement equations of the the state space model reads

$$\begin{pmatrix} \pi_t \\ \gamma_t \\ \tilde{y}_t^{n_1} \\ \vdots \\ \tilde{y}_t^{n_K} \\ \Delta v_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tilde{A}_{n_1} \\ \vdots \\ \tilde{A}_{n_K} \\ c \end{pmatrix} + \begin{pmatrix} \delta'_\pi & 0 \\ \delta'_\gamma & 0 \\ \tilde{B}'_{n_1} & 0 \\ \vdots & \vdots \\ \tilde{B}'_{n_K} & 0 \\ D' & -D' \end{pmatrix} \begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ w_t \\ \tilde{\epsilon}_{1t} \\ \vdots \\ \tilde{\epsilon}_{Kt} \\ 0 \end{pmatrix}. \tag{4.1}$$

The measurement vector on the left-hand side comprises inflation, the dividend yield of a broad-based stock index, nominal bond yields of K different maturities and the real ex-dividend return of the stock index. The data used are discussed in more detail below. The right-hand side contains the model-implied counterparts, which are functions of the states, and adds – except for inflation – measurement errors. Note that the state vector contains both the factor vector and its first lag. The latter is needed for explaining capital gains on the stock index, which are a function of ΔX_t, whereas inflation, dividend yield and bond yields depend on contemporaneous X_t only.

The elements $\tilde{A}_{n_i}$, $\tilde{B}_{n_i}$, c and D in the vector and matrix of coefficients in (4.1) are in turn functions of the deep model parameters a, $\mathcal{K}$, Σ, δ_0, δ_1, λ_0 and Λ_1, as prescribed by equations (2.18) - (2.20), (2.25) and (2.26). Hence, the measurement equations are subject to the cross-equation restrictions implied by the no-arbitrage condition.

Examining the system of measurement equations in detail, the first equation identifies the first element of the state vector as observed inflation. The second equation links the second element of the state vector to the observed dividend yield, but the two can differ in each period by a measurement error w_t.[20] This error is introduced for three reasons, outlined in the following. Noting that the role of w_t is going beyond that of a 'measurement' error in a narrow sense, we will nevertheless continue to use this terminology as is common in the econometric literature on state space models.

To begin with, w_t does in fact serve the function of a measurement error in a narrow sense. Companies that constitute our stock index (the S&P 500) are paying their dividends at different days of the year. In addition, none of these firms is paying a dividend every month. Accordingly, some scheme has to be applied to approximate that the stock index is paying out dividends in a regular fashion in each month. Thus, even if we happened to use the 'correct' model and if dividends were the only form of payout to stock holders, the measurement error would account for the arbitrariness in constructing the dividend yield

maximum likelihood (they do not have latent factors) and calibrate the remaining parameters.

[20] We do not distinguish in notation between the observed dividend yield, say γ_t^{obs} and the model counterpart γ_t but rather use the latter notation for both of them as it is always clear from the context, which one is referred to. The same holds for bond yields and stock returns.

series for the stock index.[21]

Second, the measurement error shall capture the possible wedge between the theoretical concept of the 'payout yield' of the stock and the observed dividend yields. The two can differ since dividends are not the only means by which stock investors can be made participating in profits. Most prominently, stock buy-backs are an important alternative to providing cash flow to equity holders. In fact, as shown by Boudoukh, Michaely, Richardson, and Roberts (2007), the fraction of the payout to equity holders that is due to stock buy-backs has been changing over time.

The third role of the measurement error in the dividend-yield equation is related to any form of possible misspecification of the model. More specifically, it is closely related to the fact that we do not allow for a measurement error in the last equation of the system, which relates stock returns to the state vector: we assume that realized stock returns are fully explained by the dynamics of the state variables, i.e. the two latent real-interest-rate factors and the payout yield – as it follows from the model solution. It may be reasonably argued that these factors will never account for all observed movements of monthly stock returns, especially so because they cannot capture periods of 'irrational' investment behavior, and also because we rule out rational bubbles. In fact, the approach chosen here takes a completely 'fundamental' and rational view of pricing equity. Thus, as stock returns are always perfectly matched, the conditional moments of the joint evolution of current and future dividend yields and discount factors have to align in such a way that stock returns are perfectly priced given the dynamics of the state process and given the arbitrage-free pricing relation for equity. Since factor dynamics are Markovian, the future distribution of real rates and dividend yields is completely determined by the current realization of the state vector. Moreover, since risk prices are affine functions of the factor vector, the expectation of these risk prices is also determined by current state variables. Summing up, given the model structure, a set of parameters, and realized inflation, the two latent real-rate factors and the state variable representing payout yield adjust in each period in such a way that observed stock returns are aligned with their model-implied counterparts. In this respect, the size of the measurement error w_t in the second equation indicates by how far the payout yield has to deviate from the observed dividend yield in order to 'support' the observed stock return.

A measure of how much the estimated model-implied payout yield (the second element of our state vector) has to bend away *on average* from the observed dividend yield is given by the estimated standard deviation of this measurement error. If it is close to zero, measurement errors are small on average and the dynamics of observed dividend yields are sufficient to explain the variation in observed stock returns. If it is very large, empirical dividend yields are not a very useful representative of the stochastic process representing

[21] We use the dividend series constructed by Datastream, see section 4.4 on the data.

what is considered as 'payout yield' from the viewpoint of the model. Anticipating the empirical results discussed below, it turns out that the deviation of the dividend yield from the estimated model-implied payout yield can be considered as moderate.

Further measurement errors occur in the relation of model-implied and observed zero-coupon bond yields. Again, the measurement errors account for both mis-measurement in the narrow sense (bond yields are estimated zero-coupon yields) and also for any pricing error due to model misspecification.We assume that the respective standard deviations of these bond yield measurement errors are equal across maturities. This is not uncommon in the literature on affine term structure (of interest rates) models and mainly serves to reduce the number of free parameters. At the same time, however, this approach amounts to imposing the restriction that the model's fit of bond yields is similar across maturities.

Collecting all measurement errors in a $K+1$-vector $u_t := (w_t, \tilde{\epsilon}_{1t}, \ldots, \tilde{\epsilon}_{Kt})'$, we assume that u_t is serially uncorrelated and

$$u_t \sim N(0, H), \quad H = diag(h_1^2,\, h_2^2, \ldots, h_2^2) \tag{4.2}$$

Moreover, u_t is assumed to be independent of the factor innovations η_t at all times and lags.

Finally, the transition equation of the state space model represents the dynamics of the factor vector X_t and its first lag, which is implied by (2.2),

$$\begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} \mathcal{K} & 0 \\ I & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ X_{t-2} \end{pmatrix} + \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} \eta_t. \tag{4.3}$$

4.2 Parameter restrictions

The number of parameters in the model is fairly large and not all parameters are separately identifiable. In order to reduce the number of free parameters, we will impose the following parameter restrictions on the factor dynamics (2.2):

$$\begin{pmatrix} \pi_t \\ \gamma_t \\ L_{1t} \\ L_{2t} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \mathcal{K}_{11} & 0 & 0 & 0 \\ 0 & \mathcal{K}_{22} & \mathcal{K}_{23} & \mathcal{K}_{24} \\ 0 & 0 & \mathcal{K}_{33} & 0 \\ 0 & 0 & \mathcal{K}_{43} & \mathcal{K}_{44} \end{pmatrix} \begin{pmatrix} \pi_{t-1} \\ \gamma_{t-1} \\ L_{1,t-1} \\ L_{2,t-1} \end{pmatrix} + \begin{pmatrix} \Sigma_{11} & 0 & 0 & 0 \\ 0 & \Sigma_{22} & 0 & 0 \\ 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0.001 \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \\ \eta_{4t} \end{pmatrix} \tag{4.4}$$

The real-rate factors are assumed to depend solely on their own past but not on lags of inflation or the dividend yield. The respective block of the autoregressive matrix is taken as lower-triangular, which is an innocuous assumption: as these two factors are unobservable, for each law of motion with full autoregressive matrix and full variance-covariance matrix of innovations, there is an observationally equivalent representation with lower-triangular autoregressive matrix and diagonal factor innovations. The standard deviation of factor innovations is normalized to 0.001: any re-scaling of the latent factors could be accommodated by re-sizing the respective loadings in δ_1 in (2.1).[22] Finally, the two latent factors are mean-zero processes. Thus, the unconditional expectation of the risk-free one-period rate is given by δ_0 in (2.1).

Inflation is assumed to evolve independently of the factors driving the real interest rates and the dividend yield, hence the inflation process is modelled as a simple AR(1). Together with contemporaneously independent factor innovations, this implies that we have a strict separation between the real and the nominal sphere in the model and the Fisher hypothesis will hold.[23]

For the payout yield, we allow it to be driven by its own past as well as by lagged levels of the real-rate factors. Hence, we make it an empirical issue whether real rates have forecasting power for future dividend yields.

Concerning the parameterization of the dynamics of the market prices of risk, the elements of λ_0 and Λ_1 in (2.5) are generally difficult to estimate. Moreover, there is no universally applicable set of conditions under which a certain parameter structure, i.e. a set of (zero) restrictions, guarantees identifiability. One may try to estimate the respective parameters using an iterative approach, starting with a full parametrization and subsequently setting parameters with low t-values to zero as done in Ang and Piazzesi (2003). However, there has not yet been established a 'best practice' of estimating market-price-of-risk parameters in the literature. For our model, we decided to allow only the diagonal elements of Λ_1 to be different from zero:

$$\begin{pmatrix} \lambda_{\pi,t} \\ \lambda_{\gamma,t} \\ \lambda_{L_1,t} \\ \lambda_{L_2,t} \end{pmatrix} = \begin{pmatrix} \lambda_{0,1} \\ 0 \\ \lambda_{0,3} \\ \lambda_{0,4} \end{pmatrix} + \begin{pmatrix} \Lambda_{1,11} & 0 & 0 & 0 \\ 0 & \Lambda_{1,22} & 0 & 0 \\ 0 & 0 & \Lambda_{1,33} & 0 \\ 0 & 0 & 0 & \Lambda_{1,44} \end{pmatrix} \begin{pmatrix} \pi_t \\ \gamma_t \\ L_{1t} \\ L_{2t} \end{pmatrix}.$$

We allow the time-invariant parts of the market prices of risk to be non-zero, with the exception of the market price of dividend-yield risk. The single role of $\lambda_{0,2}$ is to shift the mean of stock returns, i.e. changing $\lambda_{0,2}$ only affects c in (2.26). Hence, given all other parameters affecting the average capital gain c one could use $\lambda_{0,2}$ to shift the average capital gain to any desired value without affecting D or the pricing equations for bonds. However, using our estimation procedure outlined below, it turned out that $\lambda_{0,2}$ is difficult

[22] Normalizing the standard deviation of factor innovations to one would give rise to very small estimates of the δ_1 parameters, this is why we have chosen 1E-3 instead.

[23] This assumption is also made in Bekaert and Grenadier (2001), while d'Addona and Kind (2006) and Lettau and Wachter (2007) allow for correlated real and nominal factors or innovations.

to estimate. Accordingly, we decided to let only the other model parameters determine c.

4.3 Estimation of parameters

For the model in state space form, the Kalman filter can be used to compute the likelihood $\mathcal{L}(\psi) = \ln f(Y_1, \ldots, Y_T; \psi)$, where $Y_t = (\pi_t, \gamma_t, \tilde{y}_t^{n_1}, \ldots, \tilde{y}_t^{n_K}, \Delta v_t)$ and

$$\psi = vec(a, \mathcal{K}, \Sigma, \delta_0, \delta_1, \lambda_0, \Lambda_1, h_1, h_2)$$

is a vector containing the model parameters. However, even after imposing the restrictions expounded in the previous sub-section, the number of parameters is still large, making Maximum Likelihood estimation of the full set of unknown elements in ψ numerically burdensome. Thus, we decided to use the following two-step procedure.

In the first step, we calibrate the mean of the real interest rate δ_0 to equal the difference between the average realized nominal one-month rate and the average of realized inflation. In addition, we estimate the AR(1) process of inflation, which in the model is independent of the dynamics of the other state variables. Proceeding in this way, we obtain consistent estimates of the inflation-process parameters, but may forfeit some efficiency.[24] However, facing the trade-off between efficiency and numerical stability we opted for delegating the estimation of inflation parameters to the first step.

The second step consists of maximizing the likelihood with respect to the remaining parameters. In order to prevent getting stuck at local maxima, the corresponding optimization routine has been run from different starting vectors. Given the whole set of estimated parameters, the paths of the unobservable factors are backed out using the Kalman filter and based on these, the time series of bond and equity risk premia are obtained.

4.4 The data

Estimation is based on monthly data for the United States, spanning 26 years from January 1983 to December 2008. For inflation, we use the year-on-year log difference of the consumer price index (all urban consumers, all items). Unfortunately, there is no agreement in the literature regarding the appropriate measure of inflation in empirical asset pricing models. Since we estimate the model at monthly frequency, month-on-month inflation rates appear most appropriate from a theoretical point of view. However, as the inflation rate links the nominal and real pricing kernel, the high volatility of month-on-month inflation would lead to a counterfactual volatility in real rates, given the observed nominal yields, which would hamper the interpretation of model results. This problem

[24]This is because the parameters steering inflation dynamics (a_1, $\mathcal{K}_{11}$ and Σ_{11}) also appear in the pricing relation for nominal bonds. Hence, in a full-system estimation, the variation in bond yields is informative on the inflation parameters.

has led related literature to use year-on-year inflation rates as well.[25] For the time series of aggregate stock prices, we use end-of month values of the US S&P 500. The price index and the corresponding dividend yield are taken from Datastream.[26]

Real ex-dividend stock returns are computed as month-on-month differences of the price index, from which the inflation rate is subtracted. Nominal zero-coupon bond yields for maturities of one, two, three, five, six, seven, eight and ten years are taken from Gurkaynak, Sack, and Wright (2007).[27] The data are shown in figure 1. For estimation, all variables are expressed in monthly terms, i.e. a bond yield of 3.6% would enter as 3.6/1200=0.003. Reported results are converted back to annualized rates.

[Figure 1 about here]

5 Empirical results

5.1 Parameter estimates and fit

Parameter estimates and associated t-statistics are provided in table 1. As the estimates of the diagonal elements of $\mathcal{K}$ show, all four factors are highly persistent. In particular, the autoregressive parameter of the dividend yield $\mathcal{K}_{22}$ is very close to unity. Economically, it is plausible that the dividend yield is stationary. However, for the most part of our sample, the observed dividend yield has a clearly falling trend, hence explaining the near-unity estimate of $\mathcal{K}_{22}$.[28]

[Table 1 about here]

In analogy to econometric analyses of affine term structure (of interest rates) models, where factors often show close to I(1) dynamics, we nevertheless treat all factors as stationary. Concerning other drivers of dividend yields, we allowed $\mathcal{K}_{23}$ and $\mathcal{K}_{24}$ in (4.4) to differ from zero. In fact, the lagged real-rate factors turn out to load significantly on the

[25] See, for instance, Ang and Piazzesi (2003) and Dewachter and Lyrio (2006). As an alternative approach, one may use the more jagged month-on-month series and allow for a measurement error, which would be tantamount to using some smoothed or 'core-inflation'-type measure for pricing bonds. However, this would require some additional parameters to be estimated and the results are not expected to be hugely different. Moreover, seasonality of month-on-month inflation (or unsatisfactory seasonal adjustment) may probably cause additional problems. Finally, one may actually derive the model implications for annual inflation, but this would come at the cost of enlarging the size of the state vector.

[26] This dividend yield series is highly correlated (0.993) to the dividend-price ratios implied by Robert Shiller's dataset http://www.econ.yale.edu/ shiller/data.htm.

[27] They are downloaded from the website http://www.bilkent.edu.tr/ refet/research.html.

[28] The high persistence of dividend yields is also found in other studies. For instance, Lewellen (2004) obtains an autocorrelation coefficient of 0.991 for dividend yield over the sample 1973-2000, and 0.999 for log dividend yield.

dividend yield. Their size is small, implying a tiny effect over the short run: for instance, if the third latent factor increases such that the real interest rate is increased by hundred basis points, the expected dividend yield in the next month is affected by less than one basis point (i.e. $\mathcal{K}_{2,3}/\delta_{L1}$). For longer horizons, however, the effect becomes economically significant: using the same scenario, the expected dividend yield in five years time would increase by twelve basis points (i.e. the (2,3)-element of $\mathcal{K}^{60}$ divided by δ_{L1}). Hence, via the time profile of these effects, real rate changes can eventually have a discernible impact on the term structure of expected stock returns.

The latent factors themselves are also persistent with estimates of $\mathcal{K}_{33}$ and $\mathcal{K}_{44}$ being of a dimension well in line with those obtained in the literature for affine term structure (of interest rates) models. The estimates of the $\Lambda_{1,ii}$ signify that the factors load significantly on the respective market prices of risk. Concerning the intercepts $\lambda_{0,i}$ they are significant for the dividend yield and one real-rate factor.

The standard deviation of the measurement error of dividend yields in annualized percentage terms, i.e. $1200 \cdot h_1$, amounts to 0.19 percentage points. That is, the difference between observed and model-implied dividend yields is relatively small on average. In fact, the model-implied dividend yield, which supports observed stock returns is fairly close to its counterpart in the data: figure 2 shows the observed dividend yield together with the Kalman-filtered payout-yield factor. Although there are protracted phases of deviation, the two series show a strong overall comovement.[29]

[Figure 2 about here]

Concerning the fit of bond yields, the respective standard deviation of the measurement error is very small, as $1200 \cdot h_2$ amounts to less than seven basis points. The good fit of bond yields is confirmed in figure 3, that plots observed and estimated (Kalman-filtered) bond yields, as well as in figure 4, which compares the mean of observed yields to those implied by the filtered model-implied yields.

[Figure 3 about here]

[Figure 4 about here]

5.2 Estimated term premia

As our joint stock-bond model nests an essentially affine term structure (of interest rates) model, a plausibility check of the results is given by comparing the model-implied term premia (yield risk premia) with comparable ones obtained from a well-known affine term structure (of interest rates) model: figure 5 plots the model-implied ten-year nominal

[29]The correlation between the two series amounts to 0.98, the means are 2.58% (data) and 2.59% (model-implied), the standard deviations 1.02% and 1.03%, respectively.

term premia as defined in (2.36) together with those obtained by Kim and Wright (2005). Their premia are derived from an affine arbitrage-free three-factor term structure model. In addition to observed nominal bond yields, it employs survey data regarding the three-month Treasury Bill rate over medium and long-term horizons. From 1992 (the start of the Kim-Wright data) to 2008, the two series of estimated premia share similar dynamics (correlation coefficient of 0.8). Our premium is somewhat higher on average (1.67% vs. 1.36%) but shows less variability (0.61% vs. 0.71%). Despite their distinct comovement, the two estimates of term premia tend to diverge occasionally, in particular for the recent past: after 2005, the Kim-Wright premia showed a marked downturn to very low levels, while the premia implied by our model ranged considerably higher.

[Figure 5 about here]

5.3 The time series and cross section of estimated equity risk premia

Turning to the equity premium, figure 6 shows the estimated time series of three-month, ten-year and hundred-year premia i.e $\hat{ERP}_t^3$, $\hat{ERP}_t^{120}$ and $\hat{ERP}_t^{1200}$ defined by (2.35). It should be noted that given the estimated parameters and the filtered state vector, (2.35) can be invoked to compute the equity premium for any desired horizon. Our selection of horizons serves to illustrate the risk compensation over a very short-term horizon, a typical long-term horizon and a very long-term horizon, the latter as also employed in Lettau, Ludvigson, and Wachter (2008). For comparison, the figure also shows the equity risk premium implied by the three-stage dividend discount model described in section 3.

[Figure 6 about here]

The three time series of equity risk premia based on our model allow to broadly distinguish five phases. From 1983 until the mid-nineties, the hundred-year premium hovered around a level of about 3% with relatively small variation, while the three-month and ten-year premia showed more volatile movements around similar levels. Thereafter, equity premia displayed a distinct downward movement, reaching a long-term low before the bust of the 'dot-com boom' in early 2000.[30] With the onset of the sharp correction of this booming period, the estimated series of premia showed strong reversals. By 2003, the three-month and ten-year premia have shot up to around 5%, while the hundred-year premium reverted back to its pre 'dot com boom' level of around 3%.[31] After that, the

[30]Note that we abstain from using the term 'bubble', as from the perspective of our model, all movements in stock returns are fully rationalized implicitly by respective dividend expectations.

[31]It is interesting to compare this to the hundred-year premium shown in Lettau et al. (2008), figure 7, which is based on a consumption-based model. Also there, the premium is relatively constant first (they start in 1952) albeit on a much higher level of about 10 percent. However, their model also implies a shift to a lower premium around 1995, where it stays until 2002Q4, the end of their sample.

three-month and ten-year premium saw another downward reversal. This was brought to a halt when first signs of the financial turmoil emerged, which arose from strains in the U.S market for (subprime) mortgage-backed securities. In fact, since mid-2007, equity risk premia – especially for the short horizon – showed a remarkable increase towards the highest levels observed over the sample period. Again, the hundred-year premium was least affected, but nevertheless reached a maximum of nearly 4% by end-2008, which reflects how strongly the recent financial crisis impacted on the required risk compensation of stock market investors.

The behavior of the equity premia for the three discussed horizons suggests that longer-run measures of equity market risk compensation appear to be less volatile than those for shorter horizons. Figure 7 demonstrates this point further. It plots the sample standard deviation of all estimated time series of equity risk premia (horizon one month to hundred years) against the respective horizon, showing a monotonically decreasing pattern.

Throughout the period, for which the equity risk premium obtained from the dividend-discount model was likewise available (since 1990), the latter estimate and its counterparts implied by our model showed a strong comovement.[32] In fact, for all horizons of our estimated equity risk premium, ranging from one month towards hundred years, the correlation ranges between 0.81 and 0.92. As argued in section 3.2, the equity premium derived from the dividend-discount model may be broadly interpreted as an average across various investment horizons. In line with this interpretation is the result that the mean of of model-implied premia is closest to that of the dividend-discount model (over the shared sample period) for a horizon of about six years. Similarly, the correlation between the two series reaches a maximum at a horizon of about five years.

The comovement of the two measures of the equity risk premium (dividend-discount model vs. our model) is remarkable, given the two different approaches of estimation. The dividend-discount model makes explicit use of a forward-looking measure of future earnings, which is based on the IBES survey measure, to construct the equity risk premium. In contrast, in the affine arbitrage-free model, the equity premium for any horizon is a function of current observed state variables only. Expected stock returns result as mathematical expectations, given the estimated law of motion of the state variables and the no-arbitrage pricing relations. Hence, one interpretation of the marked comovement of the two series is that the forward-looking content of the IBES survey variable regarding the expected stock return can also be exploited from the combination of observed state variables. In the affine model, this information is extracted from the observed dividend yields and the observed term structure of interest rates. More precisely, since we actually use the *filtered* dividend yield (second state variable) and *filtered* latent real rate factors

[32]The availability of the equity risk premium obtained from the dividend discount model is limited by our access to the IBES survey data for earnings.

(third and fourth state variables), our equity risk premia are functions of current and past observed stock returns, dividend yields and bond yields.

As stressed above, the model allows to trace out at each point t in time the family of equity risk premia ERP_t^n for different investment horizons n. Figure 8 illustrates such 'term structures of equity risk premia'. In January 1999, for instance, at the climax of the 'dot-com' euphoria, the equity premium for short horizons was extremely low but then increasing towards more normal levels at the longer end. In 2002, after the stock market correction, the long end was only slightly exceeding the 1999 magnitudes, but the short end of the term structure of equity risk premia had caught up by about two percentage points. Taken together this led to a flattening of the curve. Regarding the beginning of 2008, the subprime crisis had a strong impact on required risk compensation for all investment horizons.

[Figure 8 about here]

6 Conclusion

We presented a novel approach for estimating the time series of equity risk premia. For this, we proposed a discrete-time arbitrage-free model that jointly captures stock and bond price dynamics. There is one pricing kernel that prices bonds of all maturities as well as stocks. Bond yields, bond risk premia, realized and expected stock returns, as well as equity risk premia are all affine functions of four factors: the dividend yield, the rate of inflation, and two latent factors that make up the one-period real interest rate. The model nests the class of essentially affine term structure (of interest rates) models.

With this set-up, it is possible to infer the evolution of bond premia (yield risk premia, forward premia etc) and the equity premia in a coherent framework. In addition, at each point in time, the system dynamics imply a whole 'term structure of equity risk premia'. That is, for any investment horizon n, the model provides the excess expected return of stocks over the model-implied n-period real interest rate.

Estimation is based on monthly US data from 1983 to 2008. The results make economic sense, as they comply with the intuition for prominent stock market phases such as the 'dot-com' boom phase (decreasing equity risk premia, especially over short investment horizons), the following correction (normalization of equity risk premia), as well as the onset of the 2007-08 financial turmoil (sharp increase of equity premia of all horizons). Equity risk premia for different investment horizons are positively correlated with each other. Within the estimation period, equity premia for longer horizons show less volatility than those of shorter horizons.

Furthermore, the time series of estimated equity risk premia, are strongly correlated with that based on the three-stage dividend-discount model, which gives further trust to the results. Also, the model-implied bond yield risk premia are reasonable with respect to size and dynamics as they are broadly comparable to those obtained by Kim and Wright (2005).

The comovement of the two measures of the equity risk premium (dividend-discount model vs. our model) is especially remarkable against the background that the first approach employs a forward-looking measure of future earnings (which is based on the IBES survey measure) while the affine arbitrage-free model does not. In our model, the equity premium for any horizon is a function of estimated state variables only, which are in turn filtered from current and past observations of interest rates and dividend yields. Expected stock returns result as mathematical expectations, given the estimated law of motion of the state variables and the no-arbitrage pricing relations. Hence, the results suggest that any information on expected stock returns coming from survey information can also be obtained from observed asset prices – channelled through the no-arbitrage equations of the model.

The results point to various avenues of future research. First, our specification excludes any nexus between real stock returns and inflation by assumption, i.e. through the parameter restrictions on factor dynamics. Hence, by relaxing this real-nominal orthogonality restriction, our model may contribute to the literature dealing with the impact of nominal factors on real stock prices and returns.[33] Second, instead of working with one representative stock index only, it is conceivable to apply the model to different stock portfolios in order to analyze the cross-section of equity premia in the joint stock-bond framework. Finally, there is an active literature on so-called macro-finance models of the term structure, often employing the affine framework.[34] The idea is to identify macroeconomic driving forces for bond yields and risk premia, or to use bond yield data to improve inference on structural macroeconomic relationships. Our model may be used for similar analyses of this type, allowing to trace the joint effect of macroeconomic shocks (output gap, monetary policy, natural real interest rate, etc.) on bond as well as stock market risk premia.

[33]See the survey by Sellin (2001).

[34]See, e.g., Ang and Piazzesi (2003), Ang, Dong, and Piazzesi (2007), Dewachter and Lyrio (2006), Hördahl et al. (2006), Lemke (2008), Rudebusch and Wu (2007) or Rudebusch and Wu (2008).

References

ANG, A., DONG, S., AND PIAZZESI, M. (2007). No-Arbitrage Taylor Rules. *NBER Working Paper*, 13448.

ANG, A. AND PIAZZESI, M. (2003). A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics*, 50:745–787.

BEKAERT, G. AND GRENADIER, S. R. (2001). Stock and Bond Pricing in an Affine Economy. mimeo.

BOUDOUKH, J., MICHAELY, R., RICHARDSON, M., AND ROBERTS, M. R. (2007). On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. *Journal of Finance*, 62:877–915.

CAMPBELL, J., LO, A., AND MACKINLAY, A. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

COCHRANE, J. (2001). *Asset Pricing*. Princeton University Press, Princeton et al.

— (2008). The Dog That Did Not Bark: A Defense of Return Predictability. *The Review of Financial Studies*, 21:1533–75.

D'ADDONA, S. AND KIND, A. H. (2006). International Stock-Bond Correlations in a Simple Affine Asset Pricing Model. *Journal of Banking and Finance*, 30:2747–65.

DAI, Q. AND SINGLETON, K. J. (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance*, 55:1943–1978.

DEWACHTER, H. AND LYRIO, M. (2006). Macro Factors and the Term Structure of Interest Rates. *Journal of Money, Credit, and Banking*, 38:119–140.

DUFFEE, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57:405–443.

DUFFIE, D. AND KAN, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance*, 6:379–406.

FULLER, R. J. AND HSIA, C. (1984). A Simplified Common Stock Valuation Model. *Financial Analysts Journal*, 40:49–56.

GURKAYNAK, R. S., SACK, B., AND WRIGHT, J. H. (2007). The U.S. Treasury Yield Curve: 1961 to Present. *Journal of Monetary Economics*, 54:2291–2304.

HAMILTON, J. (1994). *Time Series Analysis*. Princeton University Press, Princeton, NJ.

Harrison, J. M. and Kreps, D. (1979). Martingales and Arbitrage in Multi-Period Securities Markets. *Journal of Economic Theory*, 20:381–408.

Hördahl, P., Tristani, O., and Vestin, D. (2006). A Joint Econometric Model of Macroeconomic and Term Structure Dynamics. *Journal of Econometrics*, 131:405–444.

Kim, D. and Wright, J. (2005). An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates. *Federal Reserve Board Discussion Paper*, 33.

Lemke, W. (2006). *Term Structure Modeling and Estimation in a State Space Framework*. Springer, Berlin et al.

— (2008). An Affine Macro-Finance Term Structure Model for the Euro Area. *The North American Journal of Economics and Finance*, 19:41 – 69.

Lettau, M., Ludvigson, S. C., and Wachter, J. A. (2008). The Declining Equity Premium: What Role Does Macroeconomic Risk Play? *Review of Financial Studies*, 21:1653–87.

Lettau, M. and Wachter, J. A. (2007). The Term Structures of Equity and Interest Rates. mimeo.

Lewellen, J. (2004). Predicting Returns with Financial Ratios. *Journal of Financial Economics*, 74:209–35.

Li, L. (2002). Macroeconomic Factors and the Correlation of Stock and Bond Returns. *Yale ICF Working Paper*, (02-46).

Mamaysky, H. (2002). On the Joint Pricing of Stocks and Bonds: Theory and Evidence. *Yale ICF Working Paper*, (00-68).

Panigirtzoglou, N. and Scammell, R. (2002). Analysts' earnings forecasts and equity valuations. In *Bank of England Quarterly Bulletin, Spring 2002*, pages 59–66.

Rudebusch, G. and Wu, T. (2007). Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models. *Journal of Money, Credit and Banking*, 39:395–422.

Rudebusch, G. D. and Wu, T. (2008). A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy. *Economic Journal*, 118:906–26.

Sellin, P. (2001). Monetary Policy and the Stock Market – Theory and Empirical Evidence. *Journal of Economic Surveys*, 15:491–541.

Singleton, K. J. (2006). *Empirical Dynamic Asset Pricing*. Princeton University Press.

A Details of derivations

A.1 Arbitrage-free stock prices

We derive the solution (2.24) for arbitrage-free stock prices with coefficients defined in (2.25) and (2.26).

The derivation starts with the guess that stock prices satisfy

$$V_t = \exp[c(t-t_0) + D'X_t), \quad \text{or} \quad v_t := \ln V_t = c(t-t_0) + D'X_t \tag{A.1}$$

and then chooses D and c that makes (2.23) hold as an almost-sure identity.

Plugging the guess (A.1) into the right-hand-side of (2.23) yields

$$\begin{aligned}
& E_t \{M_{t+1}\Gamma_{t+1}V_{t+1}\} \\
= \; & E_t \{\exp[m_{t+1} + \gamma_{t+1} + v_{t+1}]\} \\
= \; & E_t \left\{\exp[-\frac{1}{2}\lambda_t'\lambda_t - \delta_0 - \delta_1' X_t - \lambda_t'\eta_{t+1} + \delta_\gamma' X_{t+1} + c(t+1-t_0) + D'X_{t+1}]\right\} \\
= \; & E_t \left\{\exp[-\frac{1}{2}\lambda_t'\lambda_t - \delta_0 + (\delta_\gamma + D)'a + c(t+1-t_0) + (\delta_\gamma + D)'\mathcal{K}X_t - \delta_1' X_t \right. \\
& \left. + \left((\delta_\gamma + D)'\Sigma - \lambda_t'\right)\eta_{t+1}]\right\}
\end{aligned}$$

This expression is of the form $E_t \exp[W_{t+1}]$, where W_{t+1} is conditionally normal. For the conditional expectation and the conditional variance we obtain

$$E_t W_{t+1} = -\frac{1}{2}\lambda_t'\lambda_t - \delta_0 + (\delta_\gamma + D)'a + c(t+1-t_0) + (\delta_\gamma + D)'\mathcal{K}X_t - \delta_1' X_t$$

and

$$Var_t W_{t+1} = \left((\delta_\gamma + D)'\Sigma - \lambda_t'\right)\left(\Sigma'(\delta_\gamma + D) - \lambda_t\right) = (\delta_\gamma + D)'\Sigma\Sigma'(\delta_\gamma + D) + \lambda_t'\lambda_t - 2(\delta_\gamma + D)'\Sigma\lambda_t$$

respectively. Using that $E_t \exp[W_{t+1}] = \exp\left[E_t W_{t+1} + \frac{1}{2} Var_t W_{t+1}\right]$, we finally get

$$\begin{aligned}
& E_t \{M_{t+1}\Gamma_{t+1}V_{t+1}\} \\
= \; & \exp\Big[-\delta_0 + (\delta_\gamma + D)'a + \frac{1}{2}(\delta_\gamma + D)'\Sigma\Sigma'(\delta_\gamma + D) - (\delta_\gamma + D)'\Sigma\lambda_0 \\
& + c\,(t+1-t_0) \\
& + \Big((\delta_\gamma + D)'\mathcal{K} - \delta_1' - (\delta_\gamma + D)'\Sigma\Lambda_1\Big)X_t\Big]
\end{aligned} \tag{A.2}$$

which completes our computation of the right-hand-side of (2.23). Using the guess (A.1), the left-hand side of (2.23) reads

$$\exp\Big[c\,(t-t_0) + D'X_t\Big].$$

In order for (2.23) to hold as an identity, the coefficients D and c have to satisfy

$$(\delta_\gamma + D)'\mathcal{K} - \delta_1' - (\delta_\gamma + D)'\Sigma\Lambda_1 = D' \tag{A.3}$$

and

$$c\,(t+1-t_0)-\delta_0+(\delta_\gamma+D)'a+\frac{1}{2}(\delta_\gamma+D)'\Sigma\Sigma'(\delta_\gamma+D)-(\delta_\gamma+D)'\Sigma\lambda_0=c\,(t-t_0) \quad \text{(A.4)}$$

respectively, for all t. Solving (A.3) for D yields (2.25), and given this solution, the expression (2.26) for c is then obtained from (A.4).

A.2 Stock returns when $\lambda_0 = 0$ and $\Lambda_1 = 0$

We derive (2.30). For $\Lambda_1 = 0$ and $\lambda_0 = 0$,

$$\begin{aligned} D' &= (\delta_\gamma'\mathcal{K}-\delta_1')(I-\mathcal{K})^{-1} \\ c &= \delta_0-(\delta_\gamma+D)'a-\frac{1}{2}(\delta_\gamma+D')\Sigma\Sigma'(\delta_\gamma+D). \end{aligned}$$

Noting that in this case $\delta_\gamma'+D'=(\delta_\gamma'-\delta_1')(I-\mathcal{K})^{-1}$ and recalling that $EX_t=(I-\mathcal{K})^{-1}a=:$ μ_X is the unconditional expectation of the stationary factor process (2.2), we have

$$c=\mu_r-\mu_\gamma-J$$

where $\mu_r := Er_t$, $\mu_\gamma = E\gamma_t$, and $J=\frac{1}{2}(\delta_\gamma+D)'\Sigma\Sigma'(\delta_\gamma+D)$. Hence, for the one-period stock return

$$\begin{aligned} & \Delta v_{t+1}+\gamma_{+1} \\ = \;& c+D'\Delta X_{t+1}+\delta_\gamma' X_{t+1} \\ = \;& c-D'X_t+(\delta_\gamma+D)'a+(\delta_\gamma+D)'\mathcal{K}X_t+(\delta_\gamma+D)'\Sigma\eta_{t+1}+\delta_\gamma'X_t-\delta_\gamma'X_t \\ = \;& c-(\delta_\gamma+D)'(I-\mathcal{K})X_t+(\delta_\gamma+D)'a+\delta_\gamma'X_t+(\delta_\gamma+D)'\Sigma\eta_{t+1} \\ = \;& \mu_r-\mu_\gamma-J-(\delta_\gamma'-\delta_1')X_t+(\delta_\gamma'-\delta_1')(I-\mathcal{K})^{-1}a+\delta_\gamma'X_t+(\delta_\gamma+D)'\Sigma\eta_{t+1} \\ = \;& \mu_r-\mu_\gamma-J-\gamma_t+r_t+\mu_\gamma-\mu_r+\gamma_t+(\delta_\gamma+D)'\Sigma\eta_{t+1} \\ = \;& r_t-J+(\delta_\gamma+D)'\Sigma\eta_{t+1} \end{aligned}$$

The second equality plugs in the law of motion for X_t, the third regroups, the fourth plugs in the expressions derived above for c and $\delta_\gamma'+D'$, the fifth uses the definition of the real short rate and the dividend yield as well as their unconditional expectations.

A.3 Multi-period stock returns

We derive $n-$period stock market returns under the assumption that any dividends paid out between period t and $t+n$ are fully reinvested in the stock (index).

In period t, the investor buys $H_t = 1$ unit of the stock at the ex-dividend price V_t. In the next period, he receives total dividends $H_t \cdot D_{t+1}$ which – following our assumption – are used for buying new stock at the new ex-dividend price V_{t+1}. Hence, the number of stocks to be bought is $\Delta H_{t+1} = \frac{H_t \cdot D_{t+1}}{V_{t+1}}$, and the new number of stocks held is $H_{t+1} =$

$H_t + \Delta H_{t+1} = H_t(1+\frac{D_{t+1}}{V_{t+1}})$. Reinvestment in the next period $t+2$ follows the same pattern, and the number of stocks held when entering period $t+3$ is $H_{t+2} = H_{t+1}(1+\frac{D_{t+2}}{V_{t+2}})$. It is straightforward to see that the number of stocks held when entering period $t+i$ is recursively obtained as

$$H_{t+i} = H_{t+i-1}\left(1 + \frac{D_{t+i}}{V_{t+i}}\right). \tag{A.5}$$

The final period of the investment horizon $t+n$ is entered with H_{t+n-1} units of the stock, then the investor obtains total dividends $H_{t+n-1} \cdot D_{t+n}$, finally he sells his stocks and obtains the revenue $H_{t+n-1} \cdot V_{t+n}$.

Thus, the overall (random) log-return of this investment strategy equals

$$\begin{aligned} r_{t+n}^{(n)} &= \ln\left(H_{t+n-1}(V_{t+n} + D_{t+n})\right) - \ln(V_t) \\ &= \ln\left(H_{t+n-1}V_{t+n}\left(1 + \frac{V_{t+n}}{D_{t+n}}\right)\right) - \ln(V_t) \end{aligned}$$

Using the definitions from the main text and $\ln(H_t) =: h_t$,

$$r_{t+n}^{(n)} = h_{t+n-1} + v_{t+n} + \gamma_{t+n} - v_t.$$

From (A.5) with $H_t = 1$, one obtains

$$h_{t+n-1} = \gamma_{t+1} + \gamma_{t+2} + \ldots + \gamma_{t+n-1}$$

Hence,

$$r_{t+n}^{(n)} = v_{t+n} - v_t + \sum_{i=1}^{n} \gamma_{t+i} \tag{A.6}$$

as had to be shown.

A.4 Expected stock returns as affine functions of factors

We derive the expressions for f_n and F_n in (2.33). The n-period return (2.32) can be written as

$$r_{t+n}^{(n)} = \frac{1}{n}\sum_{i=1}^{n}\left(\Delta v_{t+i} + \gamma_{t+i}\right).$$

With $\Delta v_{t+i} = c + D'\Delta X_{t+i}$ and $\gamma_{t+i} = \delta_\gamma' X_{t+i}$, expected stock returns can be expressed in terms of the sum of expected factors and the sum of their expected first differences:

$$E_t r_{t+n}^{(n)} = \frac{1}{n}\left(c + D'\left(\sum_{i=1}^{n} E_t \Delta X_{t+i}\right) + \delta_\gamma'\left(\sum_{i=1}^{n} E_t X_{t+i}\right)\right) \tag{A.7}$$

From the dynamics of the factor process (2.2), using the relation

$$\sum_{j=1}^{i} A^j = (I - A)^{-1}\left(I - A^{i+1}\right) \tag{A.8}$$

for a finite geometric series of a square matrix A,

$$\begin{aligned} E_t X_{t+i} &= (I-\mathcal{K})^{-1}a - (I-\mathcal{K})^{-1}\mathcal{K}^i a + \mathcal{K}^i X_t, \\ E_t \Delta X_{t+i} &= \mathcal{K}^{i-1}a + (I-\mathcal{K}^{-1})\mathcal{K}^i X_t, \end{aligned}$$

and for sums of these expectations, using (A.8) again,

$$\begin{aligned} \sum_{i=1}^{n} E_t X_{t+i} &= (I-\mathcal{K})^{-1}(nI - R(\mathcal{K},n))a + R(\mathcal{K},n)X_t \\ \sum_{i=1}^{n} E_t \Delta X_{t+i} &= R(\mathcal{K},n)\mathcal{K}^{-1}a + (I-\mathcal{K}^{-1})R(\mathcal{K},n)X_t \end{aligned}$$

with

$$R(\mathcal{K},n) = \sum_{i=1}^{n} \mathcal{K}^i = (I-\mathcal{K})^{-1}\mathcal{K}(I-\mathcal{K}^n).$$

Plugging these expressions into (A.7) and collecting terms yields the expressions for f_n and F_n in (2.33).

B Tables

Parameter	Estimate	t-value
a_1	1.117E-4	.
a_2	3.375E-06	4.39
$\mathcal{K}_{11}$	0.953	.
$\mathcal{K}_{22}$	0.999	5.10
$\mathcal{K}_{23}$	-10.084E-4	-6.96
$\mathcal{K}_{24}$	-9.268E-4	-13.21
$\mathcal{K}_{33}$	0.988	7.82
$\mathcal{K}_{43}$	-0.031	-12.14
$\mathcal{K}_{44}$	0.974	13.17
Σ_{11}	3.000E-4	.
Σ_{22}	9.208E-5	13.36
δ_0	1.976E-3	.
δ_{L1}	0.139	4.09
δ_{L2}	0.342	19.56
$\lambda_{0,1}$	-0.276	-6.18
$\lambda_{0,3}$	6.649E-5	0.01
$\lambda_{0,4}$	0.045	2.28
$\Lambda_{1,11}$	-23.883	-1.81
$\Lambda_{1,22}$	-37.878	-5.19
$\Lambda_{1,33}$	9.060	3.89
$\Lambda_{1,44}$	16.251	5.21
h_1	1.569E-4	28.96
h_2	5.101E-5	18.31

Table 1: Maximum likelihood parameter estimates and estimated asymptotic t-statistics (based on the quasi-maxium likelihood estimator for potentially misspecified models, see Hamilton (1994), section 5.8). Note that the parameters a_1, $\mathcal{K}_{11}$ and Σ_{11} of the AR(1) for inflation are estimated by OLS in the first step, and δ_0 is calibrated as the average one-month real rate in our sample (2.4% in annualized terms). For estimating the parameters $\mathcal{K}_{22}$, $\mathcal{K}_{33}$ and $\mathcal{K}_{44}$, the reparameterization $\mathcal{K}_{ii} = \psi_i^2/(1+\psi_i^2)$ has been used to guarantee that $\mathcal{K}_{ii} \in [0, 1)$. The respective t-values correspond to the auxiliary parameters ψ_i.

C Figures

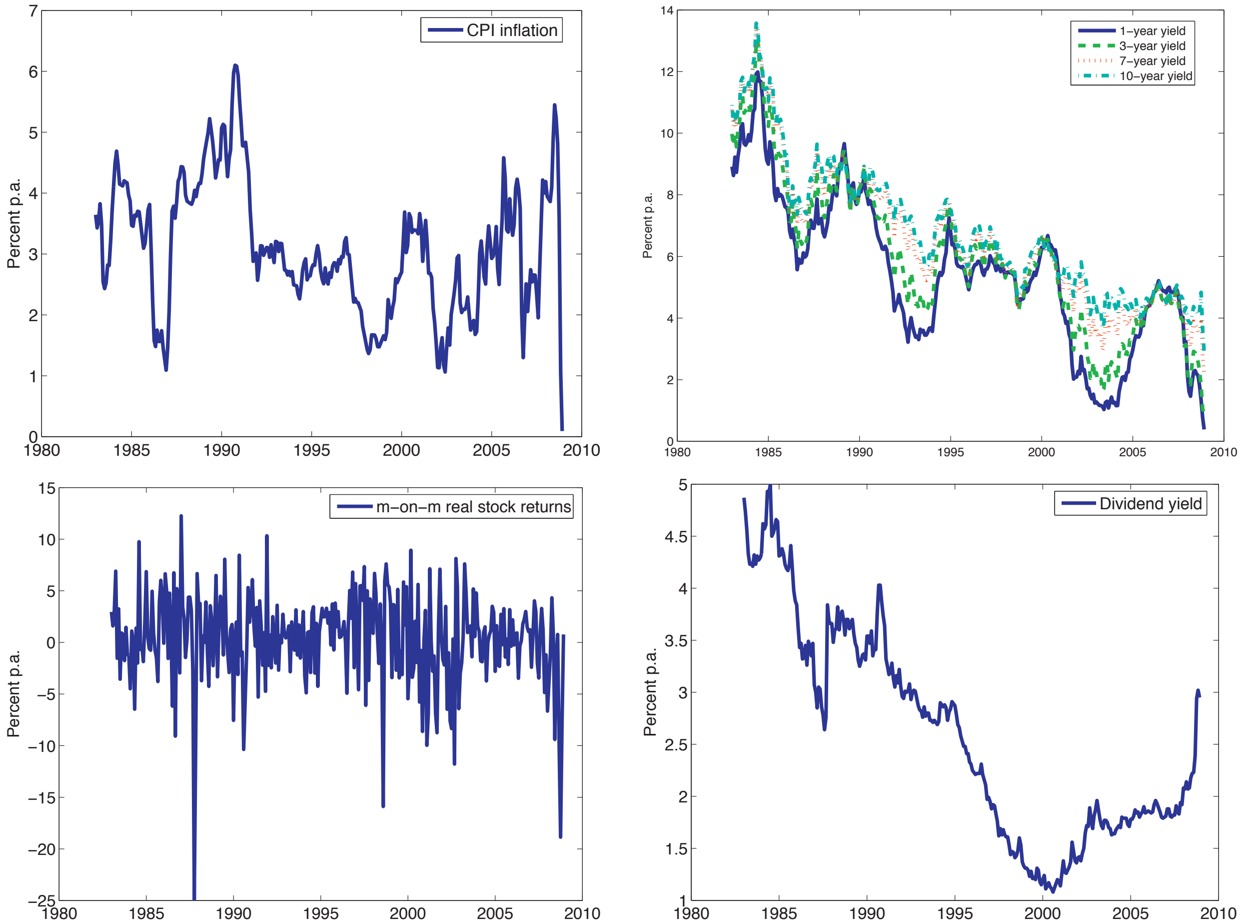

Figure 1: Monthly US Data used for estimation. CPI inflation (top left), nominal bond yields (top right), real stock returns (bottom left), dividend yield (bottom right). Jan. 1983 to Dec. 2008.

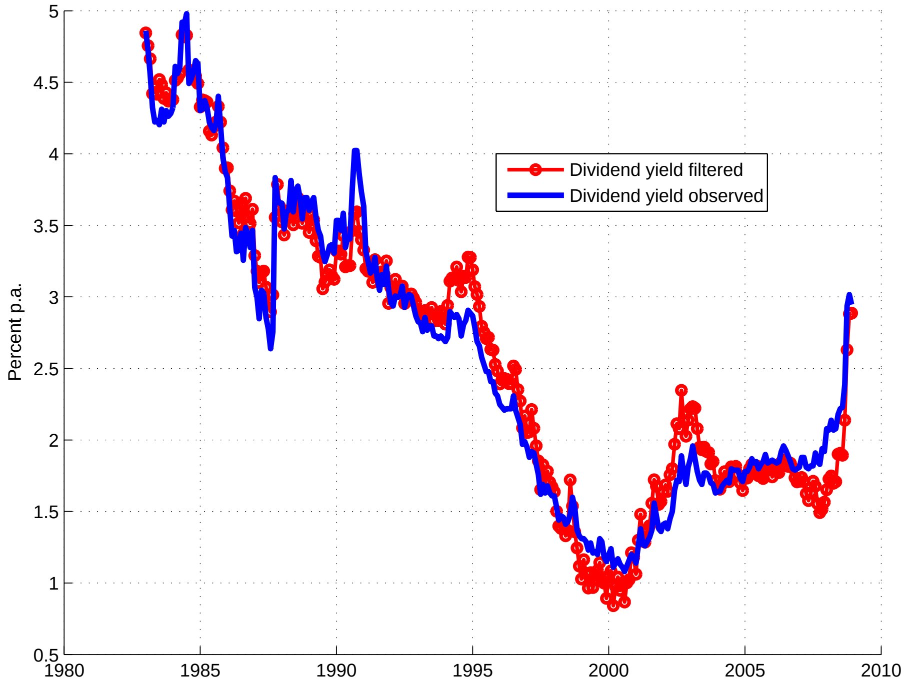

Figure 2: Dividend yield: Measured γ_t and model-implied $\hat{\gamma}_t = \delta_\gamma' X_{t|t}$, where $X_{t|t}$ denotes the filtered state vector.

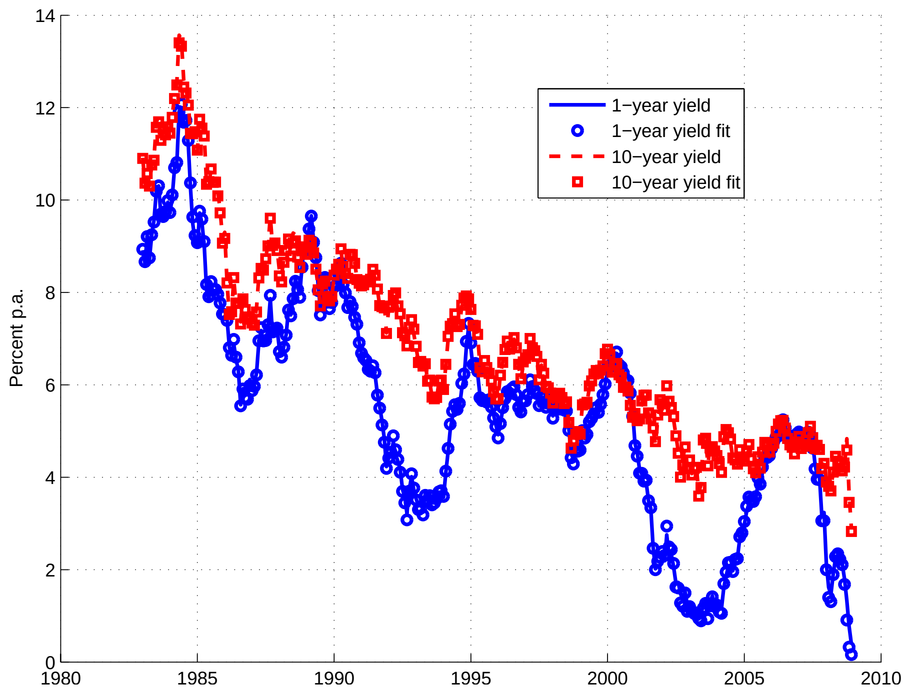

Figure 3: Observed bond yields $\tilde{y}_t^n$ for one ($n = 12$) and ten-year ($n = 120$) maturities and model-implied counterparts $\hat{\tilde{y}}_t^n = \hat{\tilde{A}}_n + \hat{\tilde{B}}_n' X_{t|t}$ based on filtered states.

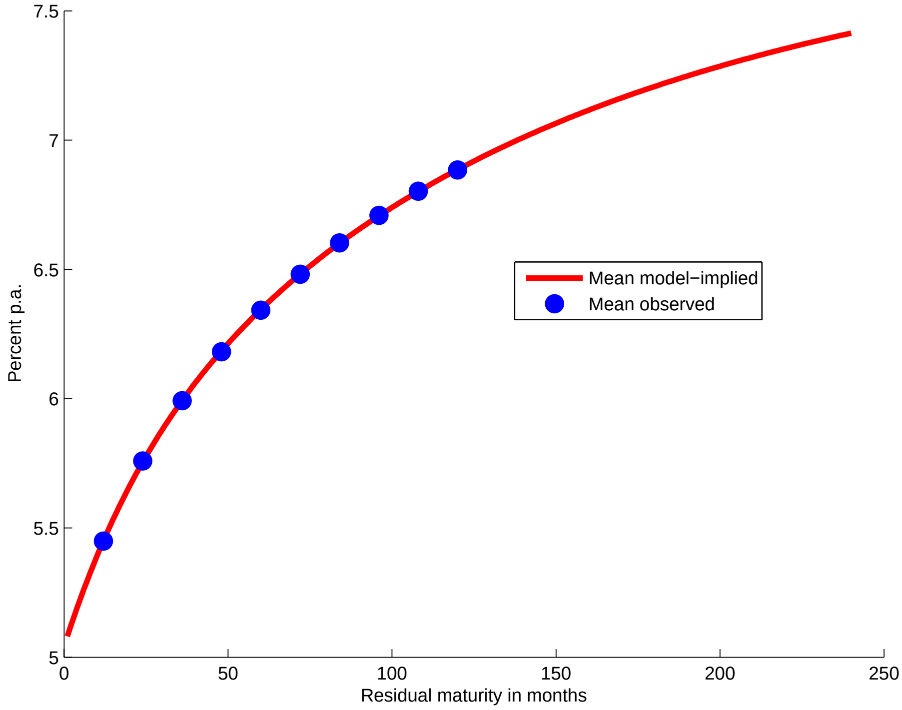

Figure 4: Average of model-implied term structure of nominal bond yields $\frac{1}{T}\sum_{t=1}^{T}\hat{\tilde{y}}_t^n$, $n = 1, \ldots, 240$ and averages of observed yields, constructed by Gurkaynak et al. (2007).

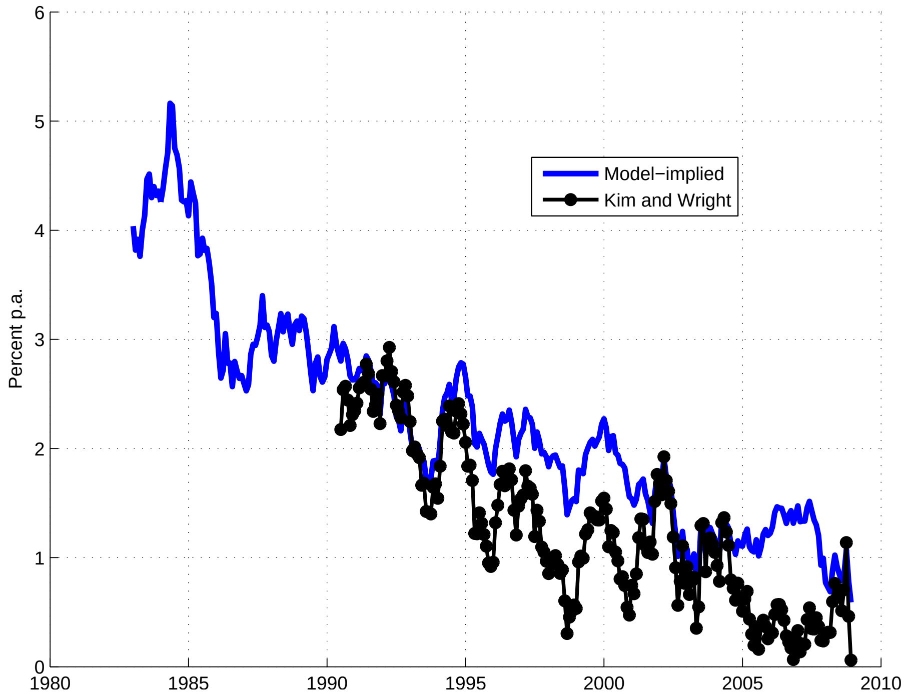

Figure 5: Ten-year nominal term premium: i) model-implied and ii) estimated by Kim and Wright (2005).

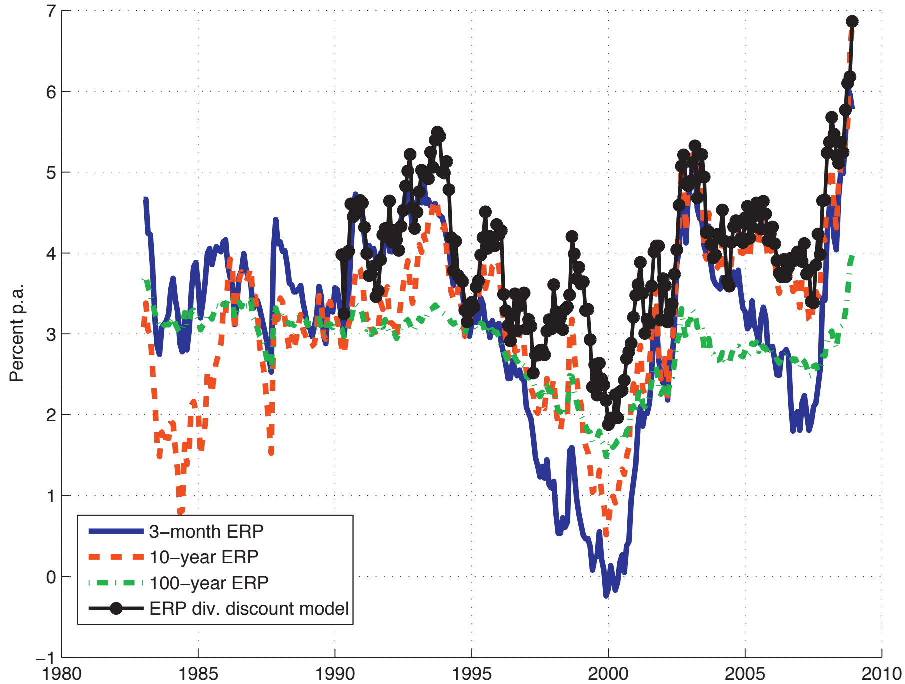

Figure 6: Model-implied equity risk premia $\widehat{ERP}_t^n$, see (2.35), for three-month, ten-year and hundred-year horizons and equity risk premium implied by three-stage dividend discount model.

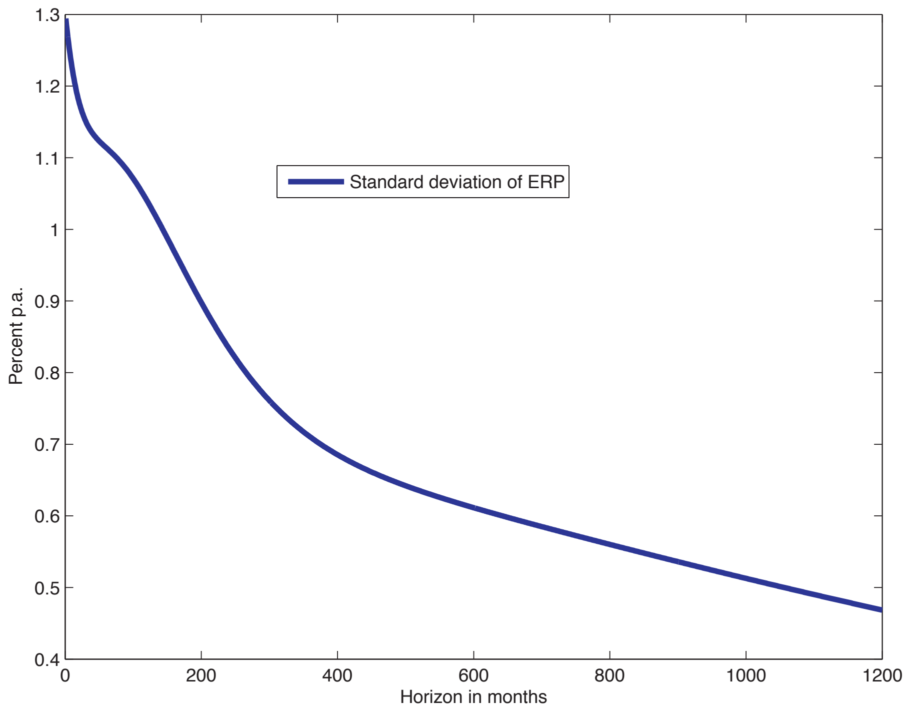

Figure 7: Standard deviation $\sqrt{Var\left(\widehat{ERP}_t^n\right)}$ of estimated time series of equity risk premia for horizons of one month to hundred years.

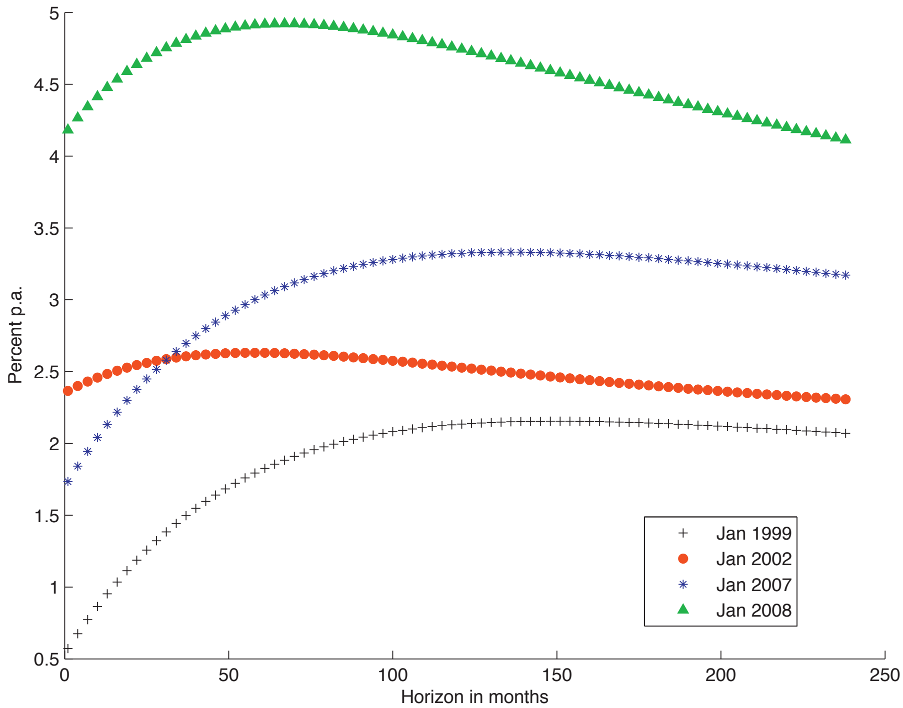

Figure 8: Equity premia for different horizons ('Term structure of ERPs') at end-January of four different years.

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.europa.eu).

989 "Modelling loans to non-financial corporations in the euro area" by C. Kok Sørensen, D. Marqués-Ibáñez and C. Rossi, January 2009

990 "Fiscal policy, housing and stock prices" by A. Afonso and R. M. Sousa, January 2009.

991 "The macroeconomic effects of fiscal policy" by A. Afonso and R. M. Sousa, January 2009.

992 "FDI and productivity convergence in central and eastern Europe: an industry-level investigation" by M. Bijsterbosch and M. Kolasa, January 2009.

993 "Has emerging Asia decoupled? An analysis of production and trade linkages using the Asian international input-output table" by G. Pula and T. A. Peltonen, January 2009.

994 "Fiscal sustainability and policy implications for the euro area" by F. Balassone, J. Cunha, G. Langenus, B. Manzke, J. Pavot, D. Prammer and P. Tommasino, January 2009.

995 "Current account benchmarks for central and eastern Europe: a desperate search?" by M. Ca' Zorzi, A. Chudik and A. Dieppe, January 2009.

996 "What drives euro area break-even inflation rates?" by M. Ciccarelli and J. A. García, January 2009.

997 "Financing obstacles and growth: an analysis for euro area non-financial corporations" by C. Coluzzi, A. Ferrando and C. Martinez-Carrascal, January 2009.

998 "Infinite-dimensional VARs and factor models" by A. Chudik and M. H. Pesaran, January 2009.

999 "Risk-adjusted forecasts of oil prices" by P. Pagano and M. Pisani, January 2009.

1000 "Wealth effects in emerging market economies" by T. A. Peltonen, R. M. Sousa and I. S. Vansteenkiste, January 2009.

1001 "Identifying the elasticity of substitution with biased technical change" by M. A. León-Ledesma, P. McAdam and A. Willman, January 2009.

1002 "Assessing portfolio credit risk changes in a sample of EU large and complex banking groups in reaction to macroeconomic shocks" by O. Castrén, T. Fitzpatrick and M. Sydow, February 2009.

1003 "Real wages over the business cycle: OECD evidence from the time and frequency domains" by J. Messina, C. Strozzi and J. Turunen, February 2009.

1004 "Characterising the inflation targeting regime in South Korea" by M. Sánchez, February 2009.

1005 "Labor market institutions and macroeconomic volatility in a panel of OECD countries" by F. Rumler and J. Scharler, February 2009.

1006 "Understanding sectoral differences in downward real wage rigidity: workforce composition, institutions, technology and competition" by P. Du Caju, C. Fuss and L. Wintr, February 2009.

1007 "Sequential bargaining in a new-Keynesian model with frictional unemployment and staggered wage negotiation" by G. de Walque, O. Pierrard, H. Sneessens and R. Wouters, February 2009.

1008 “Liquidity (risk) concepts: definitions and interactions” by K. Nikolaou, February 2009.

1009 “Optimal sticky prices under rational inattention” by B. Maćkowiak and M. Wiederholt, February 2009.

1010 “Business cycles in the euro area” by D. Giannone, M. Lenza and L. Reichlin, February 2009.

1011 “The global dimension of inflation – evidence from factor-augmented Phillips curves” by S. Eickmeier and K. Moll, February 2009.

1012 “Petrodollars and imports of oil exporting countries” by R. Beck and A. Kamps, February 2009.

1013 “Structural breaks, cointegration and the Fisher effect” by A. Beyer, A. A. Haug and B. Dewald, February 2009.

1014 “Asset prices and current account fluctuations in G7 economies” by M. Fratzscher and R. Straub, February 2009.

1015 “Inflation forecasting in the new EU Member States” by O. Arratibel, C. Kamps and N. Leiner-Killinger, February 2009.

1016 “When does lumpy factor adjustment matter for aggregate dynamics?” by S. Fahr and F. Yao, March 2009.

1017 “Optimal prediction pools” by J. Geweke and G. Amisano, March 2009.

1018 “Cross-border mergers and acquisitions: financial and institutional forces” by N. Coeurdacier, R. A. De Santis and A. Aviat, March 2009.

1019 “What drives returns to euro area housing? Evidence from a dynamic dividend-discount model” by P. Hiebert and M. Sydow, March 2009.

1020 “Opting out of the Great Inflation: German monetary policy after the break down of Bretton Woods” by A. Beyer, V. Gaspar, C. Gerberding and O. Issing, March 2009.

1021 “Rigid labour compensation and flexible employment? Firm-level evidence with regard to productivity for Belgium” by C. Fuss and L. Wintr, March 2009.

1022 “Understanding inter-industry wage structures in the euro area” by V. Genre, K. Kohn and D. Momferatou, March 2009.

1023 “Bank loan announcements and borrower stock returns: does bank origin matter?” by S. Ongena and V. Roscovan, March 2009.

1024 “Funding liquidity risk: definition and measurement” by M. Drehmann and K. Nikolaou, March 2009.

1025 “Liquidity risk premia in unsecured interbank money markets” by J. Eisenschmidt and J. Tapking, March 2009.

1026 “Do house price developments spill over across euro area countries? Evidence from a global VAR” by I. Vansteenkiste and P. Hiebert, March 2009.

1027 “Long run evidence on money growth and inflation” by L. Benati, March 2009.

1028 “Large debt financing: syndicated loans versus corporate bonds” by Y. Altunbaş, A. Kara and D. Marqués-Ibáñez, March 2009.

1029 “The role of fiscal transfers for regional economic convergence in Europe” by C. Checherita, C. Nickel and P. Rother, March 2009.

1030 "Forecast evaluation of small nested model sets" by K. Hubrich and K. D. West, March 2009.

1031 "Global roles of currencies" by C. Thimann, March 2009.

1032 "Assessing long-term fiscal developments: a new approach" by A. Afonso, L. Agnello, D. Furceri and R. Sousa, March 2009.

1033 "Fiscal competition over taxes and public inputs: theory and evidence" by S. Hauptmeier, F. Mittermaier and J. Rincke, March 2009.

1034 "The role of the United States in the global economy and its evolution over time" by S. Dées and A. Saint-Guilhem, March 2009.

1035 "The role of labor markets for euro area monetary policy" by K. Christoffel, K. Kuester and T. Linzert, March 2009.

1036 "Search in the product market and the real business cycle" by T. Y. Mathä and O. Pierrard, March 2009.

1037 "What do asset prices have to say about risk appetite and uncertainty?" by G. Bekaert, M. Hoerova and M. Scheicher, March 2009.

1038 "Are 'intrinsic inflation persistence' models structural in the sense of Lucas (1976)?" by L. Benati, March 2009.

1039 "Real Time' early warning indicators for costly asset price boom/bust cycles: a role for global liquidity" by L. Alessi and C. Detken, March 2009.

1040 "The external and domestic side of macroeconomic adjustment in China" by R. Straub and C. Thimann, March 2009.

1041 "An economic capital integrating credit and interest rate risk in the banking book" by P. Alessandri and M. Drehmann, April 2009.

1042 "The determinants of public deficit volatility" by L. Agnello and R. M. Sousa, April 2009.

1043 "Optimal monetary policy in a model of the credit channel" by F. De Fiore and O. Tristani, April 2009.

1044 "The forecasting power of international yield curve linkages" by M. Modugno and K. Nikolaou, April 2009.

1045 "The term structure of equity premia in an affine arbitrage-free model of bond and stock market dynamics" by W. Lemke and T. Werner, April 2009.