

**Regensburger  
DISKUSSIONSBEITRÄGE  
zur Wirtschaftswissenschaft**

University of Regensburg Working Papers in Business,  
Economics and Management Information Systems

**Long Memory and the Term Structure of Risk**

Peter Schotman\*, Rolf Tschernig\*\* and Jan Budek\*\*\*

January 30, 2008

Nr. 427

***JEL Classification:*** G11, C32

***Key Words:*** Long-term portfolio choice; Term structure of risk; Linear processes with fractional integration

---

\* Peter Schotman is Professor of Finance at the Limburg Institute of Financial Economics at the University of Maastricht, 6200 MD Maastricht, P.O. Box 616, Netherlands. He is also member of CEPR and NETSPAR. Phone: +31-43-388-3838, E-mail: p.schotman[at]finance.unimaas.nl

\*\* Rolf Tschernig holds the chair of Econometrics at the Department of Economics and Econometrics at the University of Regensburg, 93040 Regensburg, Germany.

Phone: +49-941-943-2737, E-mail: rolf.tschernig[at]wiwi.uni-regensburg.de

\*\*\* Jan Budek works at KPMG.

# Long Memory and the Term Structure of Risk

Peter C. Schotman<sup>a,b,c</sup>      Rolf Tschernig<sup>d</sup>      Jan Budek<sup>e</sup>

January 30, 2008

**Abstract:** This paper explores the implications of asset return predictability on long-term portfolio choice when return forecasting variables exhibit long memory. We model long memory using the class of fractionally integrated time series models. Important predictor variables for U.S. data, like the dividend-price ratio and nominal and real interest rates, are non-stationary with orders of integration around 0.8. These time series properties lead to substantial increases of the estimated long-term risk of stocks, bonds and cash compared to earlier estimates obtained from a stationary VAR. Long-term risk increases because the fluctuations in the predictor variables imply that expected returns themselves become a significant source of long-term risk. We find that results are sensitive to the specification of the prediction equation of excess stock returns. The inclusion of the short-term nominal interest rate among the predictor variables has the most profound impact. Jointly with the dividend-price ratio it has significant predictive power, but contrary to the dividend-price ratio the nominal interest rate does not induce mitigating effects through mean reversion.

**Keywords:** Long-term portfolio choice; Term structure of risk; Linear processes with fractional integration

**JEL codes:** G11, C32.

**Affiliations:** <sup>a</sup> Maastricht University, <sup>b</sup> CEPR, <sup>c</sup> NETSPAR, <sup>d</sup> University of Regensburg, <sup>e</sup> KPMG

**Correspondence address:** Peter Schotman

Limburg Institute of Financial Economics (LIFE)  
University of Maastricht, P.O. Box 616  
6200 MD Maastricht, Netherlands  
E-Mail: [P.Schotman@finance.unimaas.nl](mailto:P.Schotman@finance.unimaas.nl)  
Phone: (+31) 43 388 3838  
Fax: (+31) 43 388 4875

# 1 Introduction

A vast body of empirical research has demonstrated that expected asset returns are time-varying: short-term interest rates are used to forecast both stock and bond returns, valuation ratios such as the dividend yield or the price-earnings ratio appear to predict stock returns, and bond returns are related to the yield spread.<sup>1</sup>

Historical time series data suggest that many of the return forecasting variables display persistence. The near unit-root behavior of the predictive variables has led to an extensive literature on the properties of predictive regressions, in which the left side variable is a return series and the explanatory variable in the regression is a near-integrated series.<sup>2</sup> Inference turns out to be strongly affected when the innovations to the predictor variable are strongly correlated with unexpected returns.

An important phenomenon in this literature is mean reversion. Mean reversion is the empirical property that the long-run variance of returns is lower than the short-term variance times the investment horizon.<sup>3</sup> As a stylised fact stocks are less risky for long-term investors. Long-term investors should on average hold more stocks in their portfolio as a hedge against future changes in investment opportunities.

Although many studies have looked at testing for predictability and estimation with near-integrated series, the implications of predictability are mostly considered within a stationary vector autoregression. In this paper we study the long-run risk properties of different asset classes in models where we allow for the possibility that the prediction variables, and thus expected returns, are fractionally integrated of order  $I(d)$  ( $0 \leq d \leq 1$ ). Fractionally integrated time series models provide a flexible and still tractable linear framework with properties in between those of the stationary autoregressive  $I(0)$  models and  $I(1)$  unit root processes.<sup>4</sup>

Our analysis focuses on the term structure of risk introduced by Campbell and

---

<sup>1</sup> A brief and incomplete survey on asset return predictability includes Campbell (1991), Campbell and Shiller (1988, 1991), Ammer and Campbell (1993), Fama (1984), Fama and French (1988) and Cochrane (1997, 2007). The evidence is not undisputed: Goyal and Welch (2003, 2007) show the low out-of-sample predictability of stock returns. Continuing the debate Ang and Bekaert (2007), for example, present new favorable evidence.

<sup>2</sup> The econometrics of these unbalanced regressions have been studied in various settings. Stambaugh (1999) and Bekaert, Hodrick and Marshall (1997) derive the small sample bias for predictive regressions when the predictor is a stationary AR(1). Lewellen (2004), Campbell and Yogo (2006) and Torous, Valkanov and Yan (2004) consider the case of stocks and work under the assumption that the predictor series has an autoregressive root that is local to unity. Wachter and Warusawitharana (2007) provide a Bayesian treatment of the problem paying close attention to the prior relations between parameters close to the unit root.

<sup>3</sup> See Poterba and Summers (1988) and Fama and French (1988).

<sup>4</sup> See Robinson (2003, ch 1) and Baillie (1996) for an introduction and survey.

Viceira (2005) and its relation to the persistence of expected returns. Persistence affects the term structure in different ways. Let  $R_{t+1}$  be a one-period logarithmic return,  $q_t = \mathbf{E}_t[R_{t+1}]$  the expected log return given information up to time  $t$  and  $u_{t+1}$  the unexpected log return such that

$$R_{t+1} \equiv q_t + u_{t+1} \quad (1)$$

Adding log returns for  $k$  periods gives the cumulative log returns

$$R_{t+k}^{(k)} \equiv \sum_{j=1}^k R_{t+j} = \sum_{j=1}^k (q_{t+j-1} + u_{t+j}) \equiv q_{t+k-1}^{(k)} + u_{t+k}^{(k)} \quad (2)$$

The term structure of risk is defined as the conditional variance of cumulative returns scaled by the investment horizon,<sup>5</sup>

$$\sigma_R^2(k) \equiv \frac{1}{k} \text{Var} \left[ R_{t+k}^{(k)} - \mathbf{E}_t \left[ R_{t+k}^{(k)} \right] \right] \quad (3)$$

From the decomposition (2) we can write the term structure as the sum of three components,

$$\sigma_R^2(k) = \sigma_u^2 + 2\sigma_{uq}(k) + \sigma_q^2(k), \quad (4)$$

where  $\sigma_u^2$  is the variance of the cumulative unexpected returns  $u_{t+k}^{(k)}$  (scaled by horizon),  $\sigma_q^2(k)$  the scaled variance of cumulative expected returns  $q_{t+k-1}^{(k)}$ , and  $\sigma_{uq}(k)$  the scaled covariance. The first term,  $\sigma_u^2$ , is the usual variance of unexpected returns. By construction, since the shocks  $u_{t+1}$  are serially uncorrelated,  $\sigma_u^2$  does not depend on  $k$ . Without predictability the other two terms are zero.

Mean reversion occurs when the long-run variance is below the variance of direct shocks. i.e.  $\sigma_R^2(k) < \sigma_u^2$ . A necessary condition for this is that the covariance  $\sigma_{uq}(k)$  is negative, i.e., if unexpected returns  $u_{t+1}$  are negatively correlated with future expected returns  $q_{t+j}$ . The mean reversion effect will be stronger the higher the persistence in expected returns. When expected returns are close to a random walk, even a small shock to expected returns substantially improves the attractiveness of stocks for many periods ahead (see section 2 for a more detailed explanation).

The variance of expected returns,  $\sigma_q^2(k)$  is negligible for short horizons as the shocks to expected returns are always an order of magnitude smaller than the unexpected returns. As the variance is always positive, this term works in the opposite direction as the mean reversion effect. Both  $\sigma_{uq}(k)$  and  $\sigma_q^2(k)$  grow in absolute value with the horizon. As we will show in section 2, both diverge for any fractional process

---

<sup>5</sup> In the sequel we write the shorthand 'returns' in place of 'log returns'.

with  $d > 0$ . The limit of  $\sigma_R^2(k)$  for  $k \rightarrow \infty$  only exists in the stationary autoregressive case  $d = 0$ . For a fractional process  $\sigma_q^2(k)$  will eventually grow fastest with the horizon  $k$ . When there is predictability and some of the predictor variables are persistent, the variance of expected returns will be the dominant source of risk at long horizons. It is an empirical issue, however, which term is more important at economically relevant horizons.

For very long horizons, the fractional model perhaps overestimates the risk due to the divergence of  $\sigma_R^2(k)$ . Nevertheless, a stationary VAR model may seriously underestimate the risk, since it assumes that autocorrelations decline exponentially, and since the largest root of the system is typically underestimated in small samples.

We will study a multivariate model with three asset classes: stocks, bonds and bills. As the predictive variables for excess returns on stocks and bonds we consider the log dividend-price ratio, the nominal and real short-term interest rate, the yield spread and the credit spread. Modelling persistence requires the consideration of high-order autocorrelations, whereas the stationary first-order vector autoregressions that are widely used in the literature solely rely on low-order autocorrelations for estimation. Using semiparametric techniques we estimate the fractional integration parameter for each state variable. The fractional integration controls the long-memory or low frequency components of the time series. For the short-term dynamics we specify a first-order VAR for the fractionally differenced data. In this way we can separate the long-run and short-term dynamics in the state variables.

Predictability also implies that optimal investment strategies contain a speculative demand and a hedging demand. The hedging demand depends on the covariance of returns with the changes in future investment opportunities. Campbell and Viceira (2002) and Campbell, Chan, and Viceira (2003), in line with other authors, assume that both asset returns and return forecasting variables follow a first-order vector autoregressive process.<sup>6</sup> We derive an approximate hedging demand when the future investment opportunities follow fractional processes.

Most existing evidence for fractional integration is available for interest rates. Shea (1991), Crato and Rothman (1994) and Sun and Phillips (2004) report estimates of  $d$  of around 0.8 using a variety of estimation techniques. In a bivariate analysis of real and nominal rates Sun and Phillips (2004) find that both have the same order of integration. Backus and Zin (1993) construct the entire term structure of interest rates when the short-term interest rate is fractionally integrated, while Duan and

---

<sup>6</sup> See also Brennan and Xia (2002), Wachter (2002) and Jurek and Viceira (2006).

Jacobs (1996) show that a fractional  $I(d)$  process for the short rate can be derived in an equilibrium model.

For the stock market the possibility of fractional integration was entertained by Baillie (1996), who suggests the possibility of fractional cointegration between stock prices and dividends. An empirical test is performed by Caporale and Gil-Alana (2004) who find that stock prices and dividends fractionally cointegrate. The residuals of their cointegrating regression exhibit fractional integration with  $d > 0.5$ . Koustas and Serletis (2005) provide parametric ARFIMA estimates of  $d$  for the log dividend yield. For a long time series starting in 1871 they report estimates around 0.45 and are able to reject both the null of  $d = 0$  and  $d = 1$ . Cunado, Gil-Alana and Perez de Gracia (2005) obtain similar results for the very short sample 1994-2003.

Of particular interest is Barsky and DeLong (1993). Their proposed model for log dividends is an  $I(2)$  process and the implied log price-dividend ratio is  $I(1)$ . The motivation for their model is a coherent present value explanation for both the high volatility in stock returns and low volatility in dividend growth. Building on these insights Bansal and Yaron (2004) show that a small but persistent long-run predictable component helps in explaining the large equity premium.<sup>7</sup>

The rest of the paper is organized as follows. Section 2 discusses a univariate model for stock returns to explain the structure of the model and the effects of fractional integration. The more general multivariate model with stocks, bonds and bills is developed in section 3. For this model we discuss the term structure of risk and derive the global minimum variance portfolios for long-term investors. Section 4 discusses the estimation of the model. The empirical part starts in section 5 with a description of the data. Section 6 discusses parameter estimates, while the implied term structure of risk and the global minimum variance portfolio weights are presented in section 7. Section 8 concludes.

## 2 A Motivating Example

To motivate models with fractional integration, we start with a univariate model analysed in detail in Barberis (2000). Barberis (2000) considers an investor who can invest in either stocks or the riskfree asset. Logarithmic excess returns on equity  $y_t$

---

<sup>7</sup> Persistence in their model is measured by the first-order autoregressive parameter of expected dividend growth, which is still assumed to be stationary, but very close to one.

are linearly predictable using a state variable  $x_{t-1}$ ,

$$y_t = c + \beta x_{t-1} + u_t \quad (5)$$

Barberis (2000) further assumes that the real riskfree rate  $r_f$  is constant, so that in his model the logarithmic real return on stocks  $R_t = y_t + r_f$  differs from the excess returns by just a constant.

The state variable evolves as

$$x_t = \mu + \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j}, \quad (6)$$

where the parameters  $\beta$  and  $\theta_j$  are all scalars. The innovations  $u_t$  and  $\epsilon_t$  have variances  $\sigma_u^2$  and  $\sigma_\epsilon^2$ , respectively, and covariance  $\sigma_{u\epsilon}$ . Barberis (2000) further specialises the state dynamics to the AR(1) process

$$x_t = \mu + \alpha(x_{t-1} - \mu) + \epsilon_t, \quad (7)$$

so that  $\theta_j = \alpha^j$ .<sup>8</sup> For typical predictor variables like dividend yield, price-earnings ratio, T-Bill rate, credit spread, volatility and inflation the estimated AR parameter  $\alpha$  is always close to one and standard unit root tests usually can not reject the null hypothesis that  $x_t$  is integrated of order  $I(1)$ . The other important characteristic of the model is mean reversion. In the Barberis (2000) model a positive shock  $\epsilon_t$  to the predictive variable  $x_t$  increases next period's expected return ( $\beta > 0$ ), but has a direct negative effect on current stock returns ( $\sigma_{u\epsilon} < 0$ ). Today's bad news is partly compensated by expected higher returns in the future. For dividends this negative correlation is particularly strong as the correlation between shocks to the dividend yield and unexpected returns is close to minus one.

Predictability and mean reversion have two effects on optimal portfolios. First, investors adjust their portfolio every period to changes in expected returns. Second, as mean reversion reduces risk at longer horizons, long-term investors hold more equity on average than short-term investors. To obtain a quick view of the implications of mean reversion on the asset allocation of a long-term investor Campbell and Viceira (2005) introduce the term structure of risk  $\sigma_R^2(k)$  defined in (3). Below we analyse the effect of the time series properties of the state equation (6) on the term structure of risk.

---

<sup>8</sup> The same model consisting of (5) and (7) has been used in many other studies concerned with the predictability of stock returns and its implications. See, for example, Stambaugh (1999), Campbell and Yogo (2006), Lewellen (2004).

Even in complicated dynamic models the term structure of risk can usually be easily computed. Appendix A contains a detailed derivation of the term structure of risk for a general multivariate linear time series process. For the model consisting of the prediction equation (5) and state equation (6) the results in the appendix simplify to

$$\sigma_R^2(k) = \sigma_u^2 + 2\beta\sigma_{u\epsilon}\psi_1(k) + \beta^2\sigma_\epsilon^2\psi_2(k) \quad (8)$$

where  $\psi_1(1) = \psi_2(1) = 0$  and for  $k \geq 2$

$$\begin{aligned} \psi_1(k) &= \frac{1}{k} \sum_{\ell=0}^{k-2} \sum_{i=0}^{\ell} \theta_i \\ \psi_2(k) &= \frac{1}{k} \sum_{\ell=0}^{k-2} \left( \sum_{i=0}^{\ell} \theta_i \right)^2 \end{aligned} \quad (9)$$

The properties of the sequences  $\psi_1(k)$  and  $\psi_2(k)$  depend on the dynamic process for  $x_t$ , but not on any other parameters of the model. The term structure of risk is horizon dependent due to the dependence of  $\psi_1(k)$  and  $\psi_2(k)$  on the horizon  $k$ . For all relevant predictor variables  $\alpha$  is close to one and the impulse responses  $\theta_j$  are all positive, so that both  $\psi_1(k)$  and  $\psi_2(k)$  are monotonically increasing in  $k$ . For the AR(1) process (7), with  $\theta_i = \alpha^i$ , the sums will converge to a constant when  $0 < \alpha < 1$ .

Similar to (4), equation (8) decomposes the total risk of equity in three components. Without predictability ( $\beta = 0$ ) the last two terms in (8) are equal to zero and the entire term structure is flat at the level  $\sigma_u^2$ . This is the classical situation where portfolio choice is independent of the investment horizon.

Mean reversion may occur when  $\beta\sigma_{u\epsilon}$  is negative. The second term in (8) reflects this effect and can generate the downward slope of the term structure of risk.

The third term measures the risk from the volatility in expected returns. Since in the data the variance of innovations in expected returns is much smaller than the variance of unexpected returns ( $\beta^2\sigma_\epsilon^2 \ll \sigma_u^2$ ), this term is negligible for small  $k$ . It is always positive, though, and  $\psi_2(k)$  is increasing in  $k$ . When  $\alpha$  is very close to one, the volatility of expected returns dominates the mean reversion effect from some large  $k$  onwards, i.e. then the positive variance effect  $\beta^2\psi_2(k)\sigma_\epsilon^2$  is larger than the negative covariance effect  $2\beta\psi_1(k)\sigma_{u\epsilon}$ . Thus, for  $\alpha$  large enough, the term structure is eventually upward sloping in  $k$ .<sup>9</sup>

---

<sup>9</sup> To be precise, mean reversion reduces the long-run risk if  $\sigma_R^2(\infty) < \sigma_R^2(1)$ . To compute  $\sigma_R^2(\infty)$ , take the limits  $\psi_1(\infty) = 1/(1-\alpha)$  and  $\psi_2(\infty) = 1/(1-\alpha)^2$ . Then long-term mean reversion holds if  $2\beta\sigma_{u\epsilon}(1-\alpha) + \beta^2\sigma_\epsilon^2 < 0$ . Since  $\beta^2\sigma_\epsilon^2 > 0$ , this puts an upper limit on  $\alpha$ .



Typically, for large  $\alpha$ , the term structure is initially downward sloping, reaches a minimum and then increases again. The exact location of the turning point depends on all parameters of the model. The closer  $\alpha$  is to unity, the steeper the initial downward slope of the term structure becomes. A shock to expected returns that is expected to last longer, has more periods to compensate the immediate adverse effect. For large  $k$ , however, we find the opposite effect: the larger  $\alpha$ , the bigger the limiting risk. For the random walk ( $\alpha = 1$ ) the term structure does not converge. For  $\alpha = 1$  the volatility term  $\psi_2(k)$  becomes quadratic in  $k$  and will cause an upward sloping and diverging term structure of risk.

Figure 1 shows the effect of alternative assumptions on the time series process for  $x_t$ . For the stationary AR(1) model we use the parameter estimates from Barberis (2000). For monthly data of the dividend yield the estimated AR(1) parameter is  $\alpha = 0.9774$ . The implied term structure of risk is downward sloping and is very similar to the results shown in Campbell and Viceira (2005).<sup>10</sup> The long-run variance of equity at a 15 year horizon is about half the risk of the monthly returns. Note, however, that due to the rather large  $\alpha$  the long-run variance increases again for *very long* horizons up to the risk level of the monthly returns since  $\sigma_R^2(\infty) = 16.9\% \approx \sigma_R^2(1) = 17\%$ .

For the Barberis (2000) data the Dickey-Fuller t-statistic for a test against a unit root is -2.48, meaning that the hypothesis  $\alpha = 1$  can not be rejected at reasonable significance levels. But when  $x_t$  is a random walk the term structure of risk obtains a very different shape. In figure 1 the mean reversion is initially somewhat stronger than in the AR(1) model. After five years, however, the term structure reaches a minimum. From that point on the term structure obtains a very steep upward slope and quickly diverges. After ten years the risk is already back at the initial level. A seemingly small change to the model parameters has severe consequences for horizons that are still important for portfolio decisions.

The difference in the results is related to the functions  $\psi_1(k)$  and  $\psi_2(k)$ . These are double sums of the coefficients in the MA representation of the state variable and sensitive to the rate of decay of the  $\theta_j$  parameters. For a stationary AR(1) the term structure is of order  $O(1)$  in  $k$ , whereas under the random walk it is of order  $O(k^2)$ . Since results for the stationary AR(1) and the random walk are so different, this paper takes a closer look at the dynamics of the predictor variables. Even though we can not reject the unit root hypothesis, the behavior implied by taking the random walk model literally seems to overstate the risk of equity.

---

<sup>10</sup> See their figure 1 on page 38.

A natural candidate that is intermediate between stationary autoregressive time series processes and a random walk is a long memory process that is integrated at the fractional order  $I(d)$ . As an example, figure 1 also shows the term structure of risk for the pure  $I(d)$  process

$$(1 - L)^d(x_t - \mu) = \epsilon_t, \quad (10)$$

with long memory parameter  $d = 0.9$ . For this model the  $\theta_j$  coefficients in the MA representation (6) decline hyperbolically at the rate  $j^{d-1}$ . Consequently, the sequences  $\psi_1(k)$  and  $\psi_2(k)$  are of orders  $O(k^d)$  and  $O(k^{2d})$ , respectively, implying that the term structure is also of order  $O(k^{2d})$  for large  $k$ .<sup>11</sup> It is important to note that the term structure of risk will be upward sloping for large  $k$  even if the time series is stationary. The sequence  $\psi_2(k)$  diverges for all  $d > 0$ , whereas the returns are stationary with a finite variance for  $d < \frac{1}{2}$ .

In the figure the implied term structure of risk is almost identical to the AR(1) model up to a horizon of eight years. After that point this term structure also starts to rise, but not as steep as for the random walk model. It thus appears that the fractional model has implications that are in between the stationary AR(1) and the random walk model, but still sufficiently different from both.

In this example, the value  $d = 0.9$  is just an arbitrary number. In section 6 we provide empirical estimates for the fractional order of integration of the dividend-price ratio. We will also generalise the stylised example (5) and (6) to a model with multiple assets, multiple predictor variables and a time-varying benchmark return as to move to a framework that is closer to the model of Campbell, Chan and Viceira (2003). We allow different orders of integration for all the state variables and also more general dynamics than a pure fractional differencing model.

### 3 Multivariate Model

Let  $y_t = R_t - \iota r_{tb,t}$  be the vector of log returns on  $n$  different assets in excess of the logarithmic return of a benchmark asset  $r_{tb,t}$  where  $\iota$  denotes an  $n$  vector of ones. As in Campbell and Viceira (2005) we will use the real 3-Month T-Bill rate as the benchmark return. The state vector  $x_t$  contains  $m$  predictors. The first element of  $x_t$  will be the return on the benchmark asset. Equations (5) and (6) generalise to the

---

<sup>11</sup> See Appendix A for more details.

system

$$y_t = c + Bx_{t-1} + u_t \quad (11)$$

$$x_t = \mu + \sum_{j=0}^{\infty} \Theta_j \epsilon_{t-j}, \quad (12)$$

where the innovations  $u_t$  and  $\epsilon_t$  are now vectors of length  $n$  and  $m$  with  $(n+m)$ -dimensional covariance matrix  $\Sigma$  partitioned as

$$\Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{u\epsilon} \\ \Sigma'_{u\epsilon} & \Sigma_{\epsilon\epsilon} \end{pmatrix} \quad (13)$$

The matrices  $B$  and  $\Theta_j$  are of order  $(n \times m)$  and  $(m \times m)$ , respectively, and  $\Theta_0 = I$ .

We consider two alternative models for the dynamic behaviour of the state variables. For the first model we assume that each of the predictor variables  $x_{it}$  is fractionally integrated of order  $d_i$ . We describe the dynamics of the state vector with

$$(I - AL)\Delta(L)x_t = \epsilon_t \quad (14)$$

where the diagonal matrix  $\Delta(L)$  controls the order of integration,

$$\Delta(L) = \begin{pmatrix} (1-L)^{d_1} & & \\ & \ddots & \\ & & (1-L)^{d_m} \end{pmatrix}$$

In this model, each state variable may have a different order of fractional integration  $d_i$ , and state variables do not fractionally cointegrate among themselves. After applying the fractional filter  $\Delta(L)$ , the remaining short-term dynamics is described by a first-order VAR process with parameter matrix  $A$ . If  $d_{\max}$  is the highest order of fractional integration among the state variables predicting the  $j^{\text{th}}$  element of  $y_t$ , then the  $j^{\text{th}}$  element of  $y_t$  will be  $I(d_{\max})$ . Expected excess returns are as persistent as those state variables that enter the prediction equation (11).

For the fractional model (14) the moving-average representation (12) follows as

$$\begin{aligned} \Theta_0 &= I_m, \\ \Theta_{j+1} &= \Theta_j A + \Delta_{j+1}, \quad j \geq 0 \end{aligned} \quad (15)$$

where

$$\begin{aligned} \Delta_0 &= I_m, \\ \Delta_{j+1} &= \begin{pmatrix} \frac{j+d_1}{j+1} & & \\ & \ddots & \\ & & \frac{j+d_m}{j+1} \end{pmatrix} \Delta_j, \quad j \geq 0 \end{aligned} \quad (16)$$

Note that all  $\Delta_j$  terms for  $j > 0$  vanish when all  $d_i$ 's are equal to zero.<sup>12</sup> In case that all  $d_i$ 's are equal to zero, the model reduces to a stationary first-order VAR. When all  $d_i$ 's are equal to one, the matrices  $\Delta_j$  are all equal to the identity  $I_m$ .

For the second model we assume that the state variables are generated by a stationary VAR(2) process,

$$x_t = a + A_1 x_{t-1} + A_2 x_{t-2} + \epsilon_t \quad (17)$$

We explicitly compare the fractional model with a second-order VAR. Within a first-order VAR the parameters can only fit the long-run dynamics of the time series. A second-order VAR provides flexibility to fit both short-term and long-run dynamics as far as possible within an  $I(0)$  framework. The additional flexibility allows a better comparison with the fractionally integrated process (14). In the fractional model the state variables are transformed by the fractional filters  $(1 - L)^{d_i}$ , which take out the long-run components of the data. The remaining short-term dynamics is then captured by the first-order VAR on the filtered data. Due to the potentially different orders of integration the impulse responses of the two models considered so far can be *very* different. The  $I(0)$  assumption underlying (17) may be too restrictive as already indicated in the univariate example.

In the joint models for both  $y_t$  and  $x_t$  lagged asset returns,  $y_{t-1}$ , predict neither level returns nor state variables. We impose the restriction in order to facilitate the comparison of the fractionally integrated model with existing results. From the empirical results in Campbell and Viceira (2005) and others, this restriction is not important.

The multivariate counterpart of the term structure of risk defined in (3) is the  $((n + 1) \times (n + 1))$  covariance matrix,

$$\Sigma_{RR}^{(k)} = \frac{1}{k} \text{Var} \begin{bmatrix} R_{t+k}^{(k)} - \mathbb{E}_t \left[ R_{t+k}^{(k)} \right] \\ r_{tb,t+k}^{(k)} - \mathbb{E}_t \left[ r_{tb,t+k}^{(k)} \right] \end{bmatrix} \quad (18)$$

of the innovations of the cumulative log real returns on the benchmark asset and the  $n$  additional assets. The diagonal elements show the annualized variance of each of the asset classes. Appendix A provides detailed calculations to express the term structure as a function of the covariance parameters  $\Sigma$ , the prediction parameters  $B$  and the moving average coefficients  $\Theta_j$  for the state variables.

---

<sup>12</sup> Strictly speaking,  $(1 - L)^d$  is invertible only for  $d < 0.5$ . The actual construction of the term structure of risk exploits only a finite number of the impulse responses  $\Theta_j$ .

The construction of the term structure of risk exploits only a finite number of the cumulative impulse responses  $\Xi_k = \sum_{j=0}^k \Theta_j$ . The asymptotic behavior of the term structure of risk is similar to the univariate case. In case of a stationary VAR process, the impulse responses  $\Theta_j$  converge to zero at an exponential rate. In that case both the cumulative impulse responses and the term structure of risk converge to a constant matrix. On the contrary, if the  $i^{th}$  element of the state vector,  $x_{it}$ , is fractionally integrated with  $0 < d_i \leq 1$ , the  $(i, i)^{th}$  element of the impulse responses  $\Theta_j$  will be of order  $j^{d_i-1}$ , and the corresponding element of the partial sums  $\Xi_j$  will be of order  $j^{d_i}$ . Suppose that  $d_{\max}$  is the largest  $d$  and that all elements of  $B$  are non-zero. Then the term structure of risk diverges when  $d_{\max} > 0$ .

In a model with multiple assets, covariances are at least as important as the variances. The term structure of risk shows the potential of intertemporal risk diversification. Diversification across assets depends on the full covariance matrix. For this reason we present the global minimum variance portfolio of a long-term investor. We consider an investor who chooses a portfolio  $\alpha_t$  of the risky assets and invests the remainder fraction of wealth  $1 - \iota' \alpha_t$  in bills. Campbell and Viceira (2002, 2005) provide the following log-linear approximation to the portfolio return,

$$R_{p,t+1} = r_{tb,t+1} + \alpha_t' (y_{t+1} + \frac{1}{2} \sigma_u^2) - \frac{1}{2} \alpha_t' \Sigma_{uu} \alpha_t, \quad (19)$$

where  $\sigma_u^2$  is the vector with the diagonal elements of  $\Sigma_{uu}$ . We assume that the investor plans to hold constant proportions of his wealth in each of the asset classes for  $k$  periods. We aggregate one-period portfolio returns assuming that the investor rebalances to the initial weights at the end of each period. Starting from the one-period log return (19), aggregation to the  $k$ -period return with fixed, horizon specific, portfolio weights  $\alpha_t^{(k)}$  gives

$$R_{p,t+k}^{(k)} = r_{tb,t+k}^{(k)} + \alpha_t^{(k)'} \left( y_{t+k}^{(k)} + \frac{k}{2} \sigma_u^2 \right) - \frac{k}{2} \alpha_t^{(k)'} \Sigma_{uu} \alpha_t^{(k)} \quad (20)$$

The variance of the  $k$ -period returns follows as

$$\frac{1}{k} \text{Var}_t \left[ R_{p,t+k}^{(k)} \right] = \sigma_{rr}^{(k)2} + 2 \alpha_t^{(k)'} \sigma_{yr}^{(k)} + \alpha_t^{(k)'} \Sigma_{yy}^{(k)} \alpha_t^{(k)}, \quad (21)$$

which uses the notation  $\sigma_{rr}^{(k)2} = \text{Var}_t[r_{tb,t+k}^{(k)}]/k$  (the lower right element in the term structure covariance matrix (18)),  $\Sigma_{yy}^{(k)} = \text{Var}[y_{t+k}^{(k)}]/k$  and  $\sigma_{yr}^{(k)} = \text{Cov}[y_{t+k}^{(k)}, r_{tb,t+k}^{(k)}]/k$ . Minimising (21) with respect to  $\alpha_t^{(k)}$  we obtain the minimum variance portfolio for a horizon of  $k$  periods as

$$\alpha_H^{(k)} = -(\Sigma_{yy}^{(k)})^{-1} \sigma_{yr}^{(k)}, \quad (22)$$

The global minimum variance portfolio depends on the horizon  $k$ , but not on the state  $x_t$ . It is directly related to the term structure of risk  $\Sigma_{yy}^{(k)}$  and the term structure of covariances with the benchmark asset  $\sigma_{yr}^{(k)}$ .

For the asymptotic behavior of the portfolio weights as  $k \rightarrow \infty$  the important part is the risk due to cumulative expected excess returns. In Appendix A these terms are calculated as

$$\Sigma_{yy}^{(k)} = O(B\Psi_2(k)B') \quad (23)$$

$$\sigma_{yr}^{(k)} = O(B\Psi_2(k)h), \quad (24)$$

with  $h$  a vector that selects the element  $r_{tb,t}$  from  $x_t$  and with

$$\Psi_2(k) = \frac{1}{k} \sum_{\ell=0}^{k-2} \Xi_{\ell} \Sigma_{\epsilon\epsilon} \Xi'_{\ell},$$

a function of the cumulative impulse responses  $\Xi_{\ell}$ . To study the long-horizon properties of the hedging demands  $\alpha_H^{(k)}$  as  $k \rightarrow \infty$  first note that the  $(i, j)$ -th element of  $\Psi_2(k)$  is  $O(k^{d_i+d_j})$ , with  $d_i$  the fractional order of integration of  $x_{i,t}$ . Let  $d_{max}$  be the maximal order of integration among the  $m$  state variables and  $d_r$  the order of integration of the real T-Bill rate. Since  $r_{tb}$  is one of the state variables,  $d_r \leq d_{max}$ . The limiting behavior of  $\alpha_H^{(k)}$  follows from the properties of  $\Psi_2(k)$  and equations (23) and (24) as,

$$\begin{aligned} \Sigma_{yy}^{(k)} &\sim O(k^{2d_{max}}) \\ \sigma_{yr}^{(k)} &\sim O(k^{d_r+d_{max}}) \\ \alpha_H^{(k)} &\sim O(k^{d_r-d_{max}}) \end{aligned} \quad (25)$$

In general, when the real T-Bill does not have the highest order of integration, i.e.  $d_r < d_{max}$ , the hedging demand for stocks and bonds will go to zero for very long horizons. The minimum variance portfolio at very long horizons will then consist fully of bills. This contrasts to the usual result in Campbell and Viceira (2002) and Brennan and Xia (2002), where it is concluded that long-term bonds are the best asset class for long-term investors. Only in the special case that the real T-Bill rate has an equally high order of integration as the other assets,  $d_r = d_{max}$ , all three assets can be important in the infinite-horizon minimum variance portfolio.

So far we assumed that all state variables have predictive power for all excess returns. In some of our empirical models we will have exclusion restrictions on  $B$ . The dividend-price ratio, for example, only predicts stock returns but not bond returns. When  $B$  is subject to exclusion restrictions, the elements in  $\Sigma_{yy}^{(k)}$  and  $\sigma_{yr}^{(k)}$  can be of

different orders. Let  $d_s$  and  $d_b$  be the orders of integration of expected excess stock returns and expected excess bond returns, respectively. Depending on  $B$  and  $d_{max}$ , the orders  $d_b$  and  $d_s$  may be larger or smaller than  $d_r$ . An interesting case is a model in which the variables that predict excess returns on stocks are more persistent than the ones that predict excess bond returns and where the latter are even less persistent than the real T-Bill, leading to the inequality  $d_b < d_r < d_s$ . The configuration affects the limiting behavior of  $\alpha_H^{(k)}$ , since now

$$\begin{aligned}\Sigma_{yy}^{(k)} &\sim \begin{pmatrix} O(k^{2d_s}) & O(k^{d_s+d_b}) \\ O(k^{d_s+d_b}) & O(k^{2d_b}) \end{pmatrix} \\ \sigma_{yr}^{(k)} &\sim \begin{pmatrix} O(k^{d_s+d_r}) \\ O(k^{d_r+d_b}) \end{pmatrix}\end{aligned}\tag{26}$$

and consequently

$$\alpha_H^{(k)} \sim \begin{pmatrix} O(k^{d_r-d_s}) \\ O(k^{d_r-d_b}) \end{pmatrix}\tag{27}$$

Stocks remain unattractive and will have negligible weight at long horizons, but the weight of bonds will increase at rate  $k^{d_r-d_b}$ . It depends on the sign of the long-run correlation between bonds and T-Bills whether the weight will be positive or negative. The asymptotic properties may not be visible at economically relevant horizons though, since the divergence will be very slow if the differences between the orders of integration are small. For example, if the difference  $d_s - d_r = 0.1$ , it takes more than 14,000 periods to change the portfolio weight by a factor of 2.6.

## 4 Estimation

Estimation proceeds in two stages. First, we estimate the fractional orders of integration  $d_i$  of each of the components  $x_{it}$  of the predictor variables separately using semiparametric techniques. Using the estimated  $d_i$  we construct the fractional differences

$$w_t = \Delta(L)x_t\tag{28}$$

and estimate the parameters of the first-order VAR model

$$(I - AL)w_t = \epsilon_t\tag{29}$$

in a second stage.

Various semiparametric estimators for the long memory parameter  $d$  have been suggested: among them, the seminal log periodogram estimator of Geweke and Porter-Hudak (1983), the Gaussian semiparametric estimator (GSP) introduced in Künsch (1987) and intensively studied in Robinson and Henry (1999), and the exact Whittle likelihood estimator (EW) proposed by Shimotsu and Phillips (2005). These estimators are derived in the frequency domain and are based on the behavior of the spectral density function at frequencies close to the origin. These estimators minimize some distance between the periodogram and the spectral density function, or approximations thereof, at low frequencies represented by the first  $J$  Fourier frequencies  $\lambda_j = 2\pi j/T$ ,  $j = 1, 2, \dots, J$ .

There is only limited guidance from asymptotic theory as to which estimator to apply since the data in the current analysis may not conform to the required conditions. Violations could be caused by conditional heteroskedasticity or by extra additive noise. For example, only the GSP estimator was shown to have an asymptotic distribution that is robust to conditional heteroskedasticity (Robinson and Henry (1999)). On the other hand, the EW estimator avoids a well-known theoretical drawback of the GSP estimator given by the dependence of the asymptotic distribution on the underlying degree of long memory.<sup>13</sup> Neither the GSP, nor the EW estimator, is robust against an additive perturbation of a long memory process by white noise. Sun and Phillips (2004) show that this problem is relevant for the time series of real interest rates.

Due to the limitations of asymptotic theory we conducted a small Monte Carlo study on the performance of the estimators mentioned.<sup>14</sup> We considered data generating mechanisms that exhibit exclusively short memory or exclusively long memory or both. All those processes were generated with homoskedastic and GARCH(1,1) errors and additionally polluted by additive noise. We applied the GSP estimator to levels as well as first differences and the EW estimator. The GSP estimator based on first differences exhibits a large positive bias if applied to stationary AR(1) processes with nonnegative autoregression parameter. For the processes we considered there is no noticeable difference in performance between the GSP estimator for level data and the EW estimator.

We therefore proceed as follows. We always use the GSP estimator. If the un-

---

<sup>13</sup> The asymptotic theory for the EW estimator requires the mean of the series to be known. Shimotsu (2004) provides a modification of the EW estimator that works with unknown mean.

<sup>14</sup> The tables of the Monte Carlo are not reported here, but are available at the website of the second author.



derlying data generating mechanism is expected to be nonstationary, we apply the GSP estimator to first differences. If the underlying data generating mechanism may be stationary, we compare the estimates for levels and first differences and use the former if the latter is smaller. In this way we avoid being hit by the positive bias of the GSP estimator based on differences in case of stationary AR processes.

The choice of  $J$ , also called bandwidth selection, is important. If the number of Fourier frequencies  $J$  is chosen too large and the underlying data generating mechanism exhibits short memory, e.g. stationary AR-components, then semiparametric long memory estimators are known to be biased. On the other hand, if  $J$  is chosen too small, the estimation variance is larger than necessary. In order to obtain asymptotically valid inference,  $J$  has to grow with sample size at an appropriate rate. Henry and Robinson (1996) derive automatic bandwidth selectors for estimating homoskedastic ARFIMA processes by means of the GSP estimator. In our Monte Carlo study we found that the simple rule of selecting  $J = T^{0.5}$  performs better than various automatic bandwidth selectors since the latter selected too many frequencies even in the presence of short memory. In the empirical analysis we report three values for  $J$ , each being the smallest integer larger than  $T^\alpha$ ,  $\alpha = 0.45, 0.50, 0.55$ .

Conditional on the semiparametric estimates of  $d_i$  we use the filter  $\Delta(L)$  in (28) to construct  $w_t$ . Since the fractional filter has an infinite AR representation, in practice it has to be truncated. In order to reduce the error due to truncation, we extend the vector of state variables  $x_t$  by about two years of presample observations and use the estimated fractional differences

$$\hat{w}_t = \sum_{j=0}^{t+7} \Delta_j x_{t-j}$$

The first-order VAR model (14) for the fractional differences is then estimated by OLS. The stationary VAR(2) model (17) for the levels  $x_t$  is estimated by OLS as well.

For the unrestricted estimation of  $B$  in the prediction equations (11) we also use OLS. The limiting analysis of the portfolio weights (see (??) and (27)) has shown that the implications can differ substantially depending on which predictors are included in the equation for excess returns. As there is also considerable uncertainty about which variables have predictive power for stocks returns, and since regression analysis might overstate the evidence for predictability, we investigate the impact of model selection. Given the sample size, imposing zero restrictions on  $B$  can enhance efficiency. For this we use the method of Sequential Elimination of Regressors (SER) suggested by

Brüggemann and Lütkepohl (2001).<sup>15</sup> In each step of this iterative procedure the parameter with the smallest  $t$ -value in absolute value is eliminated if it is smaller than a threshold. This threshold is chosen such that this procedure is equivalent to parameter selection using the Schwarz model selection criterion, see Brüggemann and Lütkepohl (2001) for details. If restrictions are imposed on the prediction equations or the VAR system, FIML is used.<sup>16</sup>

Note that the OLS estimates and  $t$ -statistics have in general non-standard distributions that depend on the degrees of long memory of the regressors and the error, respectively. Although we will report  $t$ -statistics in the empirical tables, these have to be interpreted with care.<sup>17</sup>

## 5 Data

In the empirical application we consider the real excess returns on stocks ( $y_s$ ) and bonds ( $y_b$ ) over the real 3-month T-Bill rate ( $r_{tb}$ ). Other predictive variables are the log nominal T-Bill rate ( $r_{nom}$ ), the log dividend-price ratio ( $dp$ ), the credit spread ( $cr$ ), and the yield spread ( $spr$ ). We examine quarterly US data for the period 1952:Q1 – 2004:Q2. Details on data sources and data construction are in Appendix B.

The return forecasting variables have been widely used in the empirical literature.<sup>18</sup> The dividend-price ratio is an indicator of fundamental value (similar variables are price-earnings, dividend yield or consumption-wealth ratio). The credit spread is a business cycle indicator. Both are often used as predictors of excess stock returns. At least since Fama (1984) the yield spread is the traditional predictor of excess bond returns. Both the nominal interest rate as well as inflation are factors in term structure models, and thus serve as predictors of both real and nominal interest rates.

Table 1 shows summary statistics. Except for the log dividend-price ratio, the sample statistics are quarterly percentages. The important stylised facts are the high historical equity premium and the low bond premium. In terms of Sharpe ratios

---

<sup>15</sup> These calculations were conducted in JMulti, version 4.2. All other results were calculated in Ox, version 4.1 using the PcFiml package, version 1.01, and the Arfima package, version 1.04.

<sup>16</sup> Note that the prediction equations (11) represent fractional cointegration relations. To our knowledge there are no results in the literature on how the potential correlation between the noises of state variables and excess return variables biases OLS or FIML estimates in the presence of fractional cointegration.

<sup>17</sup> See Section 6 in Robinson and Marinucci (2001) for more details.

<sup>18</sup> See Avramov (2002), Cremers (2002) and Goyal and Welch (2007) for extensive references of variables commonly used for predicting stock returns.

stocks are much more attractive than bonds, and hence will dominate the speculative part of a mean-variance portfolio.

Figure 2 shows the autocorrelation functions of the five predictor series. The patterns for the nominal interest rate and the dividend-price ratio are typical for a long-memory process. Autocorrelations are large and decline very slowly. The real interest rate series looks somewhat less persistent: the autocorrelations decline faster and become even negative. The credit spread looks similar to the nominal interest rate, except that the first-order autocorrelation is much lower. The yield spread shows the least signs of persistence; its autocorrelations quickly decline as one would expect for a process with small  $d$ .

## 6 Parameter Estimates

The estimation results concern three different aspects of the model: estimates of the fractional integration parameter  $d$  of the predictor variables; the short-term dynamics of the predictor variables; and the prediction equations for excess returns on stocks and bonds.

### 6.1 Long Memory

Table 2 reports Gaussian semiparametric estimates of  $d$  for each state variable with 11, 15, and 20 Fourier frequencies, respectively. The time series of log real returns and nominal returns of T-Bills and the log dividend-price ratio were differenced once prior to estimation since these series are widely believed to be nonstationary. Since interest rate spreads are expected to be stationary the log credit and log yield spread are estimated in levels. For both the nominal and real T-Bill rate and the dividend-price ratio series the estimates of  $d$  indicate nonstationarity even though for each series the estimates vary with the choice of  $J$ . Comparing the three columns one can clearly see the influence of the number of Fourier frequencies. In the sequel we will select the estimates for  $J = 15$ . The term structure and portfolio implications remain qualitatively unchanged, however, if we would select the estimates corresponding to  $J = 20$  Fourier frequencies.

For both the nominal and the real T-Bill rates we find  $d = 0.8$ . These estimates are in line with results in the literature. Sun and Phillips (2004) apply their newly developed bivariate semiparametric estimator to real interest rates and inflation and conclude that the nominal as well as the real interest rate have about the same order

of integration which lies between 0.75 and 1. Earlier estimates, e.g. in Shea (1991) and Crato and Rothman (1994), are also often around this value.

The highest value of  $d$  is obtained for the dividend-price ratio. The point estimates of  $d$  are around 0.9 and not significantly different from  $d = 1$ . If we use very few frequencies, they are sometimes even above one. Estimates for the credit spread are found to be around 0.5 and thus at the borderline of stationarity. For the yield spread all estimates are within the stationarity region. Except for the dividend-price ratio the benchmark estimates are significantly different from both zero and one.

## 6.2 Predictability

Table 3 reports regression estimates for the prediction equations (11) for excess returns on stocks and bonds. Results for this part of the model are very similar to the general findings in the literature.<sup>19</sup> When all five predictor variables are included in the regression for future equity returns, the dividend-price ratio, the credit spread and the short-term nominal interest rate exhibit a t-statistic above two in absolute value. These three factors are also selected by the Schwarz criterion for model selection. Selecting only variables with t-statistics larger than three in absolute value, the two most important variables are the dividend-price ratio and the nominal interest rate. Both have been shown to be needed together in Ang and Bekaert (2007) and Torous, Valkanov and Yan (2004).

For comparison, because it is such a widely used specification, we also report the univariate model with the dividend-price ratio as the single predictor. It fits much worse than the model with the nominal rate and credit spread. Both the  $R^2$  of the equation and the t-statistic of the parameters are much lower than in the multivariate prediction model. Interestingly, and important later on, the coefficient of the dividend-price ratio is only half the magnitude of the  $dp$  coefficient in models where the nominal interest rate is included as well.<sup>20</sup> A similarly much poorer fit is obtained when only the nominal interest rate is included. Again we see that the parameter estimate in the univariate prediction model is much smaller than it is in the multivariate regression.

The predictable component of bond returns is related to the term spread. This is the only predictor in the multivariate regression with a t-statistic above two. The

---

<sup>19</sup> See for example Campbell and Viceira (2005) and Campbell, Chan and Viceira (2003).

<sup>20</sup> See also tables 2 and 3 in Ang and Bekaert (2007). For their quarterly post-1952 samples for four different countries they report that the dividend-price parameter increases when the nominal rate is added to the predictive regression.

coefficient on the yield spread implies that high values of the spread predict a downward move in the long-term bond yield, which is the typical result in empirical term structure models. The yield spread is also the only selected by the SER procedure based on the Schwarz criterion.

The Schwarz based specification is our benchmark model for the term structure of risk. The real T-Bill return does not appear in both prediction equations. It is still an important state variable, however, since it is the return on one of the asset classes that is available to the investor.

### 6.3 Short-term dynamics

Using the estimates  $\hat{d}_i$  for each of the state variables we have constructed the fractionally filtered data  $\sum_{j=0}^{t+7} \hat{\Delta}_j x_{t-j}$ . The filter removes most of the autocorrelations from the time series of  $r_{tb}$ ,  $r_{nom}$  and  $dp$ . Table 4 contains the results for the first-order VAR for the filtered data. The upper panel provides an estimate of the unrestricted parameter matrix  $A$ . After filtering out the persistence with a high value of  $d$ , the diagonal elements for  $r_{tb}$ ,  $r_{nom}$  and  $dp$  in  $A$  are all small. For the yield spread the diagonal element is much larger, but still far from unity. Interest rate dynamics are complicated, since all five state variables exhibit large t-statistics in the equation for the real T-Bill rate.<sup>21</sup> The credit spread looks significant in four of the five equations, the exception being the dividend-price ratio equation. The dividend-price ratio is almost a pure univariate process without any interaction with the other predictor variables. The estimated process for  $dp$  is thus very close to the stylised model used in the motivating example in section 2.

The lower panel presents a more parsimonious model based on model selection using the Schwarz criterion. This is the benchmark model for the implications.

Table 5 reports the residual covariance structure for the benchmark model of excess returns and state variables. In the residual covariance matrix the most important number is the large negative correlation of  $-0.96$  between news in the dividend-price ratio and unexpected stock returns. This is the main driver of mean reversion patterns in the term structure of equity risk. In contrast, whereas the nominal short-term interest rate is an important predictor of stock returns, the innovations to the interest rate have a very low correlation with unexpected returns. This means that the nominal interest rate does not create a mean reversion effect in equity returns.

---

<sup>21</sup> Due to the two-step estimation procedure the estimates of  $A$  may in general not show standard  $\sqrt{T}$ -asymptotics. Therefore, the t-statistics in Table 4 have to be interpreted with care.

We only report the covariance matrix for the benchmark model. The covariance estimates are only minimally different for models with different specifications for either the prediction equations for excess returns (different  $B$ 's) or models with different state variable dynamics (different  $A$  or stationary VAR). The main features are always the same.<sup>22</sup> We do not report and discuss parameter estimates for the stationary VAR(2) model. The estimation results are very similar to results reported in other studies. We will only discuss the implications of the VAR(2) in comparison to the fractional model.

## 7 Implications

We start our discussion of the implications for the term structure of risk with the multivariate model resulting from the model selection. In table 3 this corresponds to the results labeled *Schwarz* for the prediction equations for stocks and bonds. We combine these prediction equations with two different assumptions about the dynamics of the state variables. The first model is the fractionally integrated system; the other the unrestricted second-order stationary VAR. Figure 3 shows the term structure of risk for the three asset classes under the alternative assumptions.

Most striking are the results for stocks. For horizons up to two years, the fractional model and the stationary VAR are almost identical and both imply mean reversion. After the two year horizon the term structure for the stationary VAR continues its downward slope towards a long-run annual risk of 8% at a 15 year horizon similar to the results in Campbell and Viceira (2005). The fractional model, however, has very different implications beyond the minimum at the two-year horizon. The term structure rises steeply towards levels far above the initial volatility level. Since the only difference between the fractional model and the stationary VAR is in the dynamics of the state variables, the steep upward slope is evidently related to the long-memory properties of the model. The results are robust to minor changes in the dynamics of the state variables. The same patterns arises if we use the fully unrestricted  $A$  matrix instead of the restricted  $A$  from the model selection. It also does not matter if we use the estimates of  $d$  with  $m = 20$  periodogram frequencies (see third column of table 2). The results are also not due to the higher value of  $d$  for the dividend-price ratio relative to the two interest rates. We considered a model in which we lowered the long-memory parameter for the dividend-price ratio to 0.8, the same as the two

---

<sup>22</sup> When we report results for the term structure or portfolios for a particular model, we always use the complete set of estimates for that model.

interest rates, and obtained essentially the same figures.

For the other two asset classes we don't see a similar steep increase. In the fractional model the risk of bills continuously increases relative to long-term bonds. The real returns of bonds become equally risky as the long-run returns on rolling over short-term bills. For bonds the term structures implied by the stationary VAR and the fractional model remain very similar up to the six year horizon.

As the results show such a much more extreme term structure for stock returns than the univariate example in section 2 (see figure 1), we investigated in more detail what causes the steep term structure and the exceptionally large long-run risk of equity. Contrary to the robustness with respect to changes in  $A$  and  $d$ , the implied term structures appear very sensitive to the specification of the prediction equation. Most important is the interaction between the dividend-price ratio and the nominal interest rate. First consider a model with only the dividend-price ratio (see the column labeled  $DivP$  in table 3). This is the multivariate model that is most closely related to the univariate model in section 2. Parameters  $A$  and  $d$  are the same as in the benchmark model.<sup>23</sup> The top panel of figure 4 shows the term structure for this model. The steep increase for equity risk has disappeared. Mean reversion is now equally strong in the fractional model as in a stationary VAR. This is due to two effects. First, the coefficient of the dividend-price ratio in the prediction equation ( $b_{s,dp}$ ) is much smaller in this model. As a result there is considerably less volatility due to time-varying expected returns at the cost of somewhat less mean reversion. Second, the nominal interest rate and the credit spread are no longer included. Since the shocks to these two predictors are only weakly correlated with unexpected stock returns, they did not contribute to any form of mean reversion effect in the benchmark model. So setting their prediction coefficients in  $B$  equal to zero reduces the volatility of expected returns. Of the two excluded predictors the nominal interest rate has the higher  $d$  and the larger volatility. Therefore the short-term interest rate is the main driver of the incredibly steep term structure of equity in figure 3.

Thus, the steeply increasing term structure at horizons beyond two years is not due to the dividend-price ratio, even though it is the time series with the highest point estimate of  $d$ . Theoretically the term structure will still be increasing in the long run, but this is only visible when we would plot the term structure of risk over an horizon of 40 years or more.

The steep increase also disappears if the dividend-price ratio is replaced by the

---

<sup>23</sup> The residual covariance matrix  $\Sigma$  has been re-estimated, since residuals of the prediction equation for equity returns are slightly different due to the further restrictions on  $B$ .

nominal interest rate as the only prediction variable in the stock prediction equation (column *RNom* in table 3). Leaving the rest of the model as in the benchmark model (except again a minor change due to re-estimating  $\Sigma$ ), the lower panel of figure 4 shows a term structure for stocks that is much flatter and without any form of mean reversion. The explanation is that without the dividend-price ratio the predictive power of the model has almost disappeared. Even the coefficient on  $r_{nom}$  in  $B$  is much lower than in the benchmark model. To obtain interesting patterns in the term structure, predictability is required. When  $B$  is close to zero, however, as in this model, the term structure for stocks will obtain the same rate of increase as the real T-Bill, since the real stock return is the excess return plus the real T-Bill return, which is still persistent.

We conclude that the results in the benchmark model are caused by the interaction between the dividend-price ratio and the nominal T-Bill rate. When each of them is included on its own, the term structure is either mean-reverting or relatively flat. When both are included the term structure is initially mean-reverting, but rapidly rises beyond the two-year horizon. Our results are thus complementary to Ang and Bekaert (2007). They conclude, after careful econometric testing, that dividend yield and the nominal interest rate jointly produce significant predictability over short horizons. Our results show that models in which both possibly non-stationary predictors are considered jointly lead to very different implications about the long-term risk of equity.

Equally important for portfolio implications are the correlations. For the benchmark model figure 5 shows all pairwise correlations between the real returns at different horizons. The most striking difference between the implications of the fractional and stationary models is in the correlations between real returns on long-term bonds and on 3-month T-Bills. Starting above 0.3, the correlation structure resulting from the fractional model drops to as low as zero within the first year but subsequently grows to more than 0.7 at the 15-year horizon and will eventually converge to one. This implies limited diversification possibilities between bonds and bills. The explanation for the large long-run correlation comes from the structure of the prediction equation for excess bond returns. The only important predictor is the yield spread. Since the yield spread has a low  $d$ , the long-term fluctuations of real bond returns, *i.e.* excess returns plus the real T-Bill, are dominated by the real T-Bill rate. The real T-Bill rate is thus a common factor in both. This level factor determines the term structure of interest rates for long-horizon decisions. The pattern generated by



the stationary VAR estimates is completely opposite, however, falling from the same 0.3 to -0.5 at long horizons. In the stationary VAR the yield spread remains much more important for long-run movements in the term structure. A stationary VAR allows more diversification benefits than the long-memory model.

The correlation between real returns on stocks and on 3-month T-Bills is affected by long memory as well. Differences are less important here, as at least the different models agree on the sign and the order of magnitude.

Having discussed the covariance properties of the asset returns the implications for the minimum variance portfolio are straightforward. Figure 6 shows the weights of stocks, bonds and bills as a function of the investment horizon. Stocks have a small weight at all horizons. Despite their mean reversion at shorter horizons, they have by far the largest volatility. The main difference between the long-memory model and the stationary VAR(2) is in the weights for bonds and bills. In both models the long-term volatility of bonds is larger than that of bills with the difference getting smaller at longer horizons. In the long-memory model bonds and bills are strongly positively correlated and for this reason bills dominate bonds at long horizons. In the stationary model the long-term correlation between bonds and bill is negative, and therefore the weight of bonds increases with the horizon.

In the figure we have plotted the portfolio weights up to a horizon of fifteen years. From the theoretical results in (27) we know the limiting behavior as  $k \rightarrow \infty$ . In the benchmark model  $B$  is restricted such that  $d_b < d_r < d_s$  and thus the weight of stocks will go to zero, while the weight of bonds will increase without bound. Due to the positive correlation between bonds and bills, and the higher volatility of bonds, the weight of bonds will become more and more negative the longer the horizon, while the weight of bills will be more than 100%. Imposing short-sell constraints the minimum variance portfolio will be 100% in bills as it already is at the fifteen year horizon. This is opposite from the results in Campbell and Viceira (2002) who show that long-term bonds are the preferred asset class for long-horizon risk-averse investors. That conclusion thus depends on the time series properties of interest rates.

For both sets of minimum variance portfolios, implied by either the fractional or the VAR(2) model, we compute the volatility (21) under both assumptions on the dynamics of the state variables. This leads to four different estimates of the risk (two portfolios each under two dynamic models). Figure 7 shows the results. The risks displayed by solid lines are computed under the fractional dynamics of the state variables whereas dotted lines represent the risks under stationary dynamics.

The figure highlights that the difference between the two models is due to the risk implied by the state variable dynamics and thus the risk model. For the overall risk assessment of a position the choice of the risk model is much more important than the actual composition of the minimum variance portfolio. Since both portfolios agree that stocks are a minor proportion of the minimum variance portfolio, the strong rise of equity risk, which was so dominant in figure 3, does not show up in figure 7. Since the portfolio weight of bills is above 80% in any portfolio setup, the risk of the minimum variance portfolio by and large mimics the term structure of risk for real returns on T-Bills in figure 3.

## 8 Conclusion

Expected returns on stocks and bonds depend on predictor variables that exhibit long memory. For commonly used predictors as the dividend-price ratio and nominal as well as real interest rates our estimates of the order of integration of these predictors are around  $d = 0.8$ . This fractional integration has a number of implications for long-term portfolio management. First, with persistent predictors the volatility of expected returns becomes an important source of risk relative to models where expected returns are constant or subject to stationary fluctuations. This source of risk becomes relatively more important the longer the investment horizon. As a result the estimated total risk of all asset classes is much higher at long horizons in a model with fractionally integrated predictors.

Second, the persistence in interest rates leads to a strong positive correlation of the cumulative returns of long-term bonds and short-term bills. This strong correlation for long horizons is independent of predictability of excess returns as long as the real return of T-Bills exhibits long memory. Long-term bonds and short-term bills become close substitutes in the minimum variance portfolio of a long-term investor. With stationary time series for interest rates a risk-averse long-term investor will invest a large fraction of his wealth in long-term bonds. In the model with fractional integration, the minimum variance portfolio is close to 100% in bills.

Third, introducing persistence can lead to even stronger mean reversion in equity returns at short to medium horizons. Mean reversion is related to the negative covariance between the unexpected returns and future expected returns. The more persistent the predictor variable, the stronger the effect of a shock on future expected returns.

Fourth, in particular we find that the combination of the dividend-price ratio and the nominal short-term interest rate leads to much higher risk estimates for stocks compared to a stationary VAR. In models without the nominal rate we observe the usual mean reversion with a downward sloping term structure of risk.

Our results show that the term structure of risk also depends very heavily on the estimates for the prediction model of the excess returns of stocks and bonds. While OLS estimation of the prediction equations for stock and bond returns is simple, it can possibly be improved upon in terms of efficiency by alternative estimators, including system estimators such as Kim and Phillips (2001) or Chen and Hurvich (2003). Furthermore, selecting reliably relevant regressors remains an open and important issue since estimators and the corresponding t-statistics do in general have no standard asymptotics when some regressors are nonstationary and fractionally cointegrated. The resulting parameter uncertainty is not taken into account by the term structures and global minimum variance portfolios. In principle a full Bayesian approach could do this.

## Appendix A Term Structure of Risk

Below we derive the term structure of risk as the covariance matrix of the innovations of the cumulative log real returns. The covariance is constructed from the covariance matrix of cumulative excess returns, the variance of the cumulative real T-Bill rate and the covariances between excess returns and the real T-Bill rate. Starting point is the joint dynamic process for excess returns and state variables defined in (11) and (12). We repeat those below omitting the constant terms

$$y_t = Bx_{t-1} + u_t \tag{A1}$$

$$x_t = \sum_{j=0}^{\infty} \Theta_j \epsilon_{t-j} \tag{A2}$$

Substituting (A2) into the prediction equation (A1) we obtain the implied process for excess returns as

$$y_t = B \sum_{j=0}^{\infty} \Theta_j \epsilon_{t-j-1} + u_t \tag{A3}$$

Define

$$z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$$

and combine (A2) and (A3) to obtain the joint process

$$z_t = \sum_{j=0}^{\infty} C_j \nu_{t-j} \quad (\text{A4})$$

where the innovations

$$\nu_t = \begin{pmatrix} u_t \\ \epsilon_t \end{pmatrix}$$

have  $((n + m) \times (n + m))$  covariance matrix (13)

$$\Sigma = \begin{pmatrix} \Sigma_{uu} & \Sigma_{u\epsilon} \\ \Sigma'_{u\epsilon} & \Sigma_{\epsilon\epsilon} \end{pmatrix},$$

and the coefficient matrices  $C_j$  are given by

$$\begin{aligned} C_0 &= \mathbf{I}_{n+m}, \\ C_j &= \begin{pmatrix} \mathbf{0}_{n \times n} & B\Theta_{j-1} \\ \mathbf{0}_{m \times n} & \Theta_j \end{pmatrix} \quad \text{for } j > 0 \end{aligned} \quad (\text{A5})$$

For future reference we also define the cumulative impulse responses

$$\begin{aligned} D_j &= \sum_{i=0}^j C_i \\ \Xi_j &= \sum_{i=0}^j \Theta_i \end{aligned}$$

We are interested in the covariance properties of the cumulative process

$$Z_{t+k}^{(k)} = \sum_{\ell=1}^k z_{t+\ell} \quad (\text{A6})$$

for various fixed values of  $k$ . The  $\ell$ -period ahead forecast of  $z_t$  is given by

$$\mathbf{E}_t[z_{t+\ell}] = \sum_{i=\ell}^{\infty} C_i \nu_{t+\ell-i} = \sum_{i=0}^{\infty} C_{i+\ell} \nu_{t-i} \quad (\text{A7})$$

implying that the  $\ell$ -period innovation can be written as

$$z_{t+\ell} - \mathbf{E}_t[z_{t+\ell}] = \sum_{i=0}^{\ell-1} C_i \nu_{t+\ell-i} = \sum_{i=1}^{\ell} C_{\ell-i} \nu_{t+i} \quad (\text{A8})$$

Consequently, we obtain the innovation in the  $k$ -period cumulative process as

$$\begin{aligned}
Z_{t+k}^{(k)} - \mathbb{E}_t \left[ Z_{t+k}^{(k)} \right] &= \sum_{\ell=1}^k (z_{t+\ell} - \mathbb{E}_t[z_{t+\ell}]) \\
&= \sum_{\ell=1}^k \sum_{i=1}^{\ell} C_{\ell-i} \nu_{t+i} \\
&= \sum_{\ell=1}^k \left( \sum_{i=0}^{k-\ell} C_i \right) \nu_{t+\ell} \\
&= \sum_{\ell=1}^k D_{k-\ell} \nu_{t+\ell}
\end{aligned} \tag{A9}$$

From (A9) we finally obtain the scaled  $k$ -period covariance matrix

$$\begin{aligned}
V^{(k)} &= \frac{1}{k} \sum_{\ell=1}^k D_{k-\ell} \Sigma D_{k-\ell}' \\
&= \frac{1}{k} \sum_{\ell=0}^{k-1} D_{\ell} \Sigma D_{\ell}'
\end{aligned} \tag{A10}$$

The  $(n \times n)$  top-left block of  $V^{(k)}$  contains the term structure of risk  $\Sigma_{yy}^{(k)}$  of excess returns, while the first element of the lower right block contains the variance  $\sigma_r^{(k)2}$  of the real T-Bill rate. The  $(n \times m)$  top-right block contains the covariances  $\sigma_{yr}^{(k)}$  with the real T-Bill. We can write these matrices explicitly by using the partitioning of the  $C_i$  matrices in (A5) and  $\Sigma$  in (13). Since  $D_0 = C_0 = I_{m+n}$  and also  $\Xi_0 = I_m$ , we have

$$D_0 \Sigma D_0' = \begin{pmatrix} \Sigma_{uu} & \Sigma_{u\epsilon} \Xi_0' \\ \Xi_0 \Sigma_{u\epsilon}' & \Xi_0 \Sigma_{\epsilon\epsilon} \Xi_0' \end{pmatrix} \tag{A11}$$

For  $j > 0$  we find

$$D_j = C_0 + \sum_{i=1}^j C_i = \begin{pmatrix} \mathbf{I}_n & B \Xi_{j-1} \\ \mathbf{0}_{m \times n} & \Xi_j \end{pmatrix} \tag{A12}$$

and thus

$$\begin{aligned}
D_j \Sigma D_j' &= \begin{pmatrix} \mathbf{I}_n & B \Xi_{j-1} \\ \mathbf{0}_{m \times n} & \Xi_j \end{pmatrix} \begin{pmatrix} \Sigma_{uu} & \Sigma_{u\epsilon} \\ \Sigma_{u\epsilon}' & \Sigma_{\epsilon\epsilon} \end{pmatrix} \begin{pmatrix} \mathbf{I}_n & \mathbf{0}_{n \times m} \\ \Xi_{j-1}' B' & \Xi_j' \end{pmatrix} \\
&= \begin{pmatrix} \Sigma_{uu} + B \Xi_{j-1} \Sigma_{u\epsilon}' + \Sigma_{u\epsilon} \Xi_{j-1}' B' + B \Xi_{j-1} \Sigma_{\epsilon\epsilon} \Xi_{j-1}' B' & \Sigma_{u\epsilon} \Xi_j' + B \Xi_{j-1} \Sigma_{\epsilon\epsilon} \Xi_j' \\ \Xi_j \Sigma_{u\epsilon}' + \Xi_j \Sigma_{\epsilon\epsilon} \Xi_{j-1}' B' & \Xi_j \Sigma_{\epsilon\epsilon} \Xi_j' \end{pmatrix}
\end{aligned} \tag{A13}$$

Summing the terms  $D_j \Sigma D_j'$  and scaling by  $k$  we finally obtain

$$\begin{aligned}\Sigma_{yy}^{(k)} &= \Sigma_{uu} + \Sigma_{u\epsilon} \Psi_1(k)' B' + B \Psi_1(k) \Sigma'_{u\epsilon} + B \Psi_2(k) B' \\ \sigma_{yr}^{(k)} &= \Sigma_{u\epsilon} \Psi_3(k)' h + B \Psi_4(k)' h \\ \sigma_{rr}^{(k)2} &= \frac{k+1}{k} h' \Psi_2(k+1) h\end{aligned}\tag{A14}$$

where  $\Psi_i(1) = 0$  ( $i = 1, \dots, 4$ ) and for  $k > 1$

$$\begin{aligned}\Psi_1(k) &= \frac{1}{k} \sum_{\ell=0}^{k-2} \Xi_{\ell} \\ \Psi_2(k) &= \frac{1}{k} \sum_{\ell=0}^{k-2} \Xi_{\ell} \Sigma_{\epsilon\epsilon} \Xi'_{\ell} \\ \Psi_3(k) &= \Psi_1(k) + \frac{1}{k} \Xi_{k-1} \\ \Psi_4(k) &= \Psi_2(k) + \frac{1}{k} \sum_{\ell=0}^{k-2} \Theta_{\ell+1} \Sigma_{\epsilon\epsilon} \Xi'_{\ell},\end{aligned}$$

The elements in  $\Sigma_{yy}^{(k)}$  contain the term structure of excess returns relative to the benchmark real T-Bill rate. To obtain the term structure of real returns we add the real T-Bill back and use

$$\Sigma_{RR}^{(k)} = \sigma_{rr}^{(k)2} \begin{pmatrix} \iota' & \iota \\ \iota' & 1 \end{pmatrix} + \begin{pmatrix} \Sigma_{yy}^{(k)} + \sigma_{yr}^{(k)} \iota' + \iota \sigma_{yr}^{(k)'} & \sigma_{yr}^{(k)} \\ \sigma_{yr}^{(k)'} & 0 \end{pmatrix}\tag{A15}$$

The diagonal elements of  $\Sigma_{RR}^{(k)}$  are shown in the term structure graphs.

For the univariate analysis with a constant benchmark return  $r_t = r_f$  in section 2,  $\sigma_{rr}^{(k)2}$  and  $\sigma_{yr}^{(k)}$  are obviously zero, and the expressions for  $\Psi_1(k)$  and  $\Psi_2(k)$  simplify to

$$\psi_1(k) = \frac{1}{k} \sum_{\ell=0}^{k-2} \xi_{\ell} \quad k \geq 2\tag{A16}$$

$$\psi_2(k) = \frac{1}{k} \sum_{\ell=0}^{k-2} \xi_{\ell}^2\tag{A17}$$

$$\xi_{\ell} = \sum_{i=0}^{\ell} \theta_i\tag{A18}$$

with  $\psi_1(1) = \psi_2(1) = 0$ . For ease of interpretation the definition of the univariate  $\psi_2(k)$  differs slightly from its multivariate counterpart  $\Psi_2(k)$  by not including  $\sigma_{\epsilon}^2$ .

To study the order of convergence of  $\psi_1(k)$  and  $\psi_2(k)$  we use the fact that if

$$f_k = \sum_{j=1}^k j^a$$

and

$$g_k = \frac{f_k}{k^{a+1}}$$

then for  $a > -1$ ,

$$g_k = \sum_{j=1}^k \left(\frac{j}{k}\right)^a \frac{1}{k} \rightarrow \int_0^1 x^a dx = \frac{1}{a+1} \quad (\text{A19})$$

Therefore,  $f_k \sim O(k^{a+1})$  if  $a > -1$ .

Applying the result to the orders of the impulse responses starting from  $\theta_j \sim O(j^{d-1})$ , we have

$$\begin{aligned} \xi_k &\sim O(k^d), \\ \psi_1(k) &\sim O(k^d), \\ \psi_2(k) &\sim O(k^{2d}), \\ \sum_{j=1}^k \theta_j^2 &\sim O(k^{2d-1}), \end{aligned} \quad (\text{A20})$$

Cumulative impulse responses and the term structure sequences  $\psi_1(k)$  and  $\psi_2(k)$  do not converge for any  $d > 0$ . On the other hand, the sum of squared impulse responses converges to a constant for  $d \leq \frac{1}{2}$ . We thus have that for  $0 < d < \frac{1}{2}$  the unconditional variance of a series exists, but the term structure of risk diverges.

## Appendix B Data sources and timing

Raw data are obtained from different sources:

**Nominal return of 3-month T-Bills** is from the FRED database of the St. Louis Fed (<http://research.stlouisfed.org/fred2/>), series *DTB3*. The quarterly data are extracted from the daily data as the interest rate at the last trading day of the quarter. The interest rate for period  $t + 1$  is known at the end of period  $t$  and denoted as *shortint<sub>t</sub>*.

**10-year constant maturity yield** is also from FRED, series *DGS10*, last day of the quarter, denoted as *longint<sub>t</sub>*.

**Corporate Bond Yield** is from FRED, series *BAA*, extracted from the monthly data as the last month of the quarter, denoted as *corp\_baa<sub>t</sub>*.

**Consumer Price Index** is taken from FRED, series *CPIAUCNS*, not seasonally adjusted. Quarterly data are obtained from the monthly series and denoted as *cpi<sub>t</sub>*.

**Stock price index** is the S&P 500, obtained from *Yahoo!Finance*. Data are the closing prices at the end of each quarter, denoted as  $sp500_t$ .

**Dividends** are from the monthly "Irrational Exuberance" data of Robert Shiller ([www.econ.yale.edu/~shiller/](http://www.econ.yale.edu/~shiller/)). They have been converted to quarterly dividends on the S&P 500 as the dividend paid until the end of period  $t$ , denoted  $dividend_t$ .

To obtain the data for the model estimation we have performed the following transformations.

**Nominal T-Bill** is defined as

$$r_{nom,t} = 100 \times \ln(1 + shortint_t/400), \quad (\text{B1})$$

and is denoted in percent per quarter.

**Excess bond returns** are constructed from the long-term interest rate. We use the log-linear approximation described in Campbell, Lo, and MacKinlay (1997, ch 10) and construct the 10-year excess bond return series as

$$y_{b,t+1} = \frac{1}{4} \mathcal{Y}_{t+1} - D_t(\mathcal{Y}_{t+1} - \mathcal{Y}_t) - r_{nom,t} \quad (\text{B2})$$

where the log bond yield is defined as

$$\mathcal{Y}_t = 100 \times \ln(1 + longint_t/100),$$

and the approximate bond duration is given by

$$D_t = \frac{1 - (1 + longint_t/100)^{-10}}{1 - (1 + longint_t/100)^{-1}}$$

We approximate the yield on a 9 year and 3 quarters bond with the yield on a 10-year bond.

**Excess stock returns** use the S&P 500 prices and dividends as

$$y_{s,t+1} = \ln \left( \frac{sp500_{t+1} + dividend_{t+1}}{sp500_t} \right) - r_{nom,t} \quad (\text{B3})$$

**Real T-Bill rate** is the nominal rate minus a quarter of the annual inflation,

$$r_{tb,t} = r_{nom,t-1} - \frac{1}{4} (\ln cpi_t - \ln cpi_{t-4}) \quad (\text{B4})$$

This is the ex-post real return computed at the end of period  $t$  using the nominal return that is already known at the end of the previous period.



**Dividend-price ratio** is defined in logarithmic form as the annual dividend divided by the latest price as

$$dp_t = \ln \left( \frac{\text{dividend}_t + \text{dividend}_{t-1} + \text{dividend}_{t-2} + \text{dividend}_{t-3}}{sp500_t} \right) \quad (\text{B5})$$

**Credit spread** is the difference in yield between corporate bonds and 10-year government bonds

$$cr_t = \frac{1}{4} \ln \left( \frac{1 + \text{longint}_t/100}{1 + \text{corp\_baa}_t/100} \right) \quad (\text{B6})$$

**Yield spread** is the difference between the yield on 10-year government bonds and the 3-month nominal T-Bill rate

$$spr_t = \frac{1}{4} \ln \left( \frac{1 + \text{longint}_t/100}{1 + \text{shortint}_t/100} \right) \quad (\text{B7})$$

## References

- AMMER, J. AND J.Y. CAMPBELL (1993), What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns, *Journal of Finance*, **48**, 3–37.
- ANG, A. AND G. BEKAERT (2007), Stock Return Predictability: Is it There?, *Review of Financial Studies*, **20**, 651–707.
- AVRAMOV, D. (2002), Stock Return Predictability and Model Uncertainty, *Journal of Financial Economics*, **64**, 423–458.
- BACKUS, D.K. AND S.E. ZIN (1993), Long-Memory Inflation Uncertainty: Evidence from the Term Structure of Interest Rates, *Journal of Money Credit and Banking*, **25**, 681–700.
- BAILLIE, R.T. (1996), Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics*, **73**, 5–59.
- BANSAL, R. AND A. YARON (2004), Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance*, **59**, 1481–1509.
- BARBERIS, N. (2000), Investing for the Long Run when Returns are Predictable, *Journal of Finance*, **55**, 389–406.
- BARSKY, R.B. AND J.B. DELONG (1993), Why Does the Stock Market Fluctuate?, *Quarterly Journal of Economics*, **108**, 291–311.
- BEKAERT, G., R.J. HODRICK, AND D.A. MARSHALL (1997), On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates, *Journal of Financial Economics*, **44**, 309–348.
- BRENNAN, M.J. AND Y. XIA (2002), Dynamic Asset Allocation under Inflation, *Journal of Finance*, **57**, 1201–1238.
- BRÜGGEMANN, R. AND H. LÜTKEPOHL (2001), Lag Selection in Subset VAR models with an Application to a U.S. Monetary System, in L.Knüppel R.Friedmann and H.Lütkepohl, (eds.), *Econometric Studies: A Festschrift in Honour of Joachim Frohn*, LIT Verlag, 107–128.

- CAMPBELL, J.Y. (1991), A Variance Decomposition for Stock Returns, *Economic Journal*, **101**, 157–179.
- CAMPBELL, J.Y., Y.L. CHAN, AND L.M. VICEIRA (2003), A Multivariate Model for Strategic Asset Allocation, *Journal of Financial Economics*, **67**, 41–80.
- CAMPBELL, J.Y., A.W. LO, AND A.C. MACKINLAY (1997), *The Econometrics of Financial Markets*, Princeton University Press.
- CAMPBELL, J.Y. AND R.J. SHILLER (1988), Stock Prices, Earnings and Expected Dividends, *Journal of Finance*, **43**, 661–676.
- CAMPBELL, J.Y. AND R.J. SHILLER (1991), Yield Spreads and Interest Rate Movements: A Bird’s Eye View, *Review of Economic Studies*, **58**, 495–514.
- CAMPBELL, J.Y. AND L.M. VICEIRA (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press.
- CAMPBELL, J.Y. AND L.M. VICEIRA (2005), The Term Structure of the Risk-Return Tradeoff, *Financial Analysts Journal*, **61**, 34–44.
- CAMPBELL, J.Y. AND M. YOGO (2006), Efficient Tests of Stock Return Predictability, *Journal of Financial Economics*, **81**, 27–60.
- CAPORALE, G. M. AND L.A. GIL-ALANA (2004), Fractional Cointegration and Tests of Present Value Models, *Review of Financial Economics*, **13**, 245–258.
- CHEN, W.W. AND C.M. HURVICH (2003), Semiparametric Estimation of Multivariate Fractional Cointegration, *Journal of the American Statistical Association*, **98**, 629–642.
- COCHRANE, J.H. (1997), Where is the Market Going? Uncertain Facts and Novel Theories, *Economic Perspectives*, **21**, 3–37, Federal Reserve Bank of Chicago.
- COCHRANE, J.H. (2007), The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies*, **20**, 1–43.
- CRATO, N. AND P. ROTHMAN (1994), Fractional Integration Analysis of Long-run Behavior for US Macroeconomic Time Series, *Economic Letters*, **45**, 287–291.
- CREMERS, K.J.M. (2002), Stock Return Predictability: A Bayesian Model Selection Perspective, *Review of Financial Studies*, **15**, 1223–1249.
- CUNADO, J., L.A. GIL-ALANA, AND F. PEREZ DE GRACIA (2005), A Test for Rational Bubbles in the NASDAQ Stock Index: A Fractionally Integrated Approach, *Journal of Banking and Finance*, **29**, 2633–2654.
- DUAN, J. AND K. JACOBS (1996), A Simple Long-memory Equilibrium Interest Rate Model, *Economic Letters*, **53**, 317–321.
- FAMA, E.F. (1984), Term Premiums in Bond Returns, *Journal of Financial Economics*, **13**, 529–546.
- FAMA, E.F. AND K.R. FRENCH (1988), Permanent and Transitory Components of Stock Prices, *Journal of Political Economy*, **96**, 246–273.
- FAMA, E.F. AND K.R. FRENCH (1989), Business Conditions and the Expected Returns on Stocks and bonds, *Journal of Financial Economics*, **25**, 23–49.
- GEWEKE, J. AND S. PORTER-HUDAK (1983), The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis*, **4**, 221–238.
- GOYAL, A. AND I. WELCH (2003), Predicting the Equity Premium with Dividend Ratios, *Management Science*, **49**, 639–654.
- GOYAL, A. AND I. WELCH (2007), A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies*, forthcoming.

- HENRY, M. AND P.M. ROBINSON (1996), Bandwidth Choice in Gaussian Semiparametric Estimation of Long Range Dependence, in *Athens Conference on Applied Probability and Time Series*, volume 115 of *Lecture Notes in Statistics*, Springer, New York, 220–232.
- JUREK, J.W. AND L.M. VICEIRA (2006), Optimal Value and Growth Tilts in Long-Horizon Portfolios, Harvard Business School Working Paper.
- KIM, C.S. AND P.C.B. PHILLIPS (2001), Fully Modified Estimation of Fractional cointegration Models, Yale Technical Report.
- KOUSTAS, Z. AND A. SERLETIS (2005), Rational Bubbles or Persistent Deviations from Market Fundamentals?, **29**, 2523–2539.
- KÜNSCH, H. (1987), Statistical Aspects of Self-Similar Processes, in Y. V. Prohorov and V. V. Sazonov, (eds.), *Proceedings of the First World Congress of the Bernoulli Society*, VNU Science Press, 67–74.
- LEWELLEN, J. (2004), Predicting Returns with Financial Ratios, *Journal of Financial Economics*, **74**, 209–235.
- POTERBA, J.M. AND L.H. SUMMERS (1988), Mean Reversion in Stock Prices, *Journal of Financial Economics*, **22**, 27–59.
- ROBINSON, P.M. (2003), *Time Series with Long Memory*, Oxford University Press.
- ROBINSON, P.M. AND D. MARINUCCI (2001), Narrow-Band Analysis of Nonstationary Processes, *Annals of Statistics*, **29**, 942–986.
- ROBINSON, P.M. AND M. HENRY (1999), Long and Short Memory Conditional Heteroskedasticity in Estimating the Memory Parameter of Levels, *Econometric Theory*, **15**, 299–336.
- SHEA, G.S. (1991), Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest Rate Term Structure, *Empirical Economics*, **16**, 287–312.
- SHIMOTSU, K. (2004), Exact Local Whittle Estimation of Fractional Integration with Unknown Mean and Time Trend, Working paper, Queens University, Kingston.
- SHIMOTSU, K. AND P.C.B. PHILLIPS (2005), Exact Local Whittle Estimation of Fractional Integration, *American Statistician*, **33**, 1890–1933.
- STAMBAUGH, R.F. (1999), Predictive Regressions, *Journal of Financial Economics*, **54**, 375–421.
- SUN, Y. AND P.C.B. PHILIPS (2004), Understanding the Fisher Equation, *Journal of Applied Econometrics*, **19**, 869–886.
- TOROUS, W., R. VALKANOV, AND S. YAN (2004), On Predicting Stock Returns with Nearly Integrated Explanatory Variables, *Journal of Business*, **77**, 937–966.
- WACHTER, J. (2002), Optimal Consumption and Portfolio Allocation under Mean-Reverting Returns: An Exact Solution for Complete Markets, *Journal of Financial and Quantitative Analysis*, **37**, 63–91.
- WACHTER, J. AND M. WARUSAWITHARANA (2007), Predictable returns and Asset Allocation: Should a Skeptical Investor Time the Market?, NBER Working Paper 13165.

Table 1: Descriptive statistics, 1952 Q1 – 2004 Q2

	avg	std.	min	max	skew	kurt
Excess log stock returns	1.44	7.91	-30.76	18.97	-0.92	4.83
Excess log bond returns	0.18	3.91	-13.86	13.61	0.35	4.47
Log real return of 3-month T-Bills	0.33	0.50	-1.39	1.91	-0.07	3.92
Log nominal return of 3-month T-Bills	1.26	0.70	0.16	3.52	0.96	4.17
Log dividend-price ratio	-3.45	0.40	-4.49	-2.77	-0.83	3.38
Log credit spread	0.39	0.17	0.09	0.17	0.50	2.77
Log yield spread	0.35	0.28	-0.45	1.00	0.08	2.73

*Notes:* All variables are quarterly. The log returns and log spreads are given in percentages. "Skew" and "Kurt" are the scaled skewness and kurtosis of the time series.

Table 2: Gaussian semiparametric estimates of  $d$ 

	pre-filter	Number of frequencies ( $J$ )		
		12	15	20
Log real return of 3-month T-Bills	diff	0.677	<b>0.796</b>	0.727
Log nominal return of 3-month T-Bills	diff	0.733	<b>0.800</b>	0.848
Log dividend-price ratio	diff	1.081	<b>0.877</b>	0.924
Log credit spread	diff	0.201	0.261	0.316
	level	0.474	<b>0.525</b>	0.553
Log yield spread	diff	0.026	0.308	0.438
	level	0.147	<b>0.320</b>	0.455
(Standard error)		(0.144)	(0.129)	(0.112)

*Notes:* Results are obtained using the Gaussian semiparametric estimator (GSP) in Robinson and Henry (1999). The number of frequencies  $J$  is determined by the smallest integer larger than  $T^\alpha$ ,  $\alpha \in \{0.45, 0.50, 0.55\}$ . Prior to estimation all series are differenced once before. The  $d$  values are reported for the level series and are obtained by adding one. These results are indicated by *diff* in the second column. If the estimated  $d$  is less than 0.5, the value of  $d$  is re-estimated directly on the levels, without first differencing. These results are indicated by *level*. Standard errors are reported from the series of first differences. Numbers in **boldface** are the estimates that are used in the further analysis.

Table 3: Prediction equations for excess log stock and bond returns

	<i>Unrestricted</i>		<i>Schwarz</i>		<i>DivP</i>	<i>RNom</i>
	Stocks	Bonds	Stocks	Bonds	Stocks	Stocks
constant	0.222 (0.060)	-0.033 (0.030)	0.207 (0.054)	-0.009 (0.004)	0.102 (0.048)	0.021 (0.016)
$r_{tb,t}$	1.849 (1.180)	0.909 (0.585)				
$r_{nom,t}$	-3.378 (1.010)	-0.622 (0.501)	-3.033 (0.835)			-1.812 (0.786)
$dp_t$	0.060 (0.017)	-0.006 (0.008)	0.055 (0.015)		0.025 (0.014)	
$cr_t$	9.546 (4.187)	-2.684 (2.078)	9.133 (3.438)			
$spr_t$	-0.679 (2.418)	3.887 (1.200)		3.087 (0.945)		
$R^2$	0.089	0.081	0.077	0.049	0.016	0.024

*Notes:* The unrestricted system is estimated by OLS. The restricted systems are estimated using FIML using the restricted specification for the excess bond returns. Standard errors in parenthesis. Regressors are a constant ( $c$ ), the log real T-Bill rate ( $r_{tb,t}$ ), the log nominal T-Bill rate ( $r_{nom,t}$ ), the dividend-price ratio ( $dp_t$ ), the credit spread ( $cr_t$ ) and the yield spread ( $spr_t$ ).

Table 4: Fractional VAR(1) model

<b>Unrestricted parameter estimates</b>						
	$r_{tb,t}$	$r_{nom,t}$	$dp_t$	$cr_t$	$spr_t$	$R^2$
$r_{tb,t+1}$	0.200 (0.047)	1.132 (0.085)	-0.005 (0.001)	0.608 (0.169)	0.459 (0.087)	0.565
$r_{nom,t+1}$	-0.136 (0.066)	-0.240 (0.119)	-0.002 (0.002)	-0.816 (0.237)	-0.210 (0.122)	0.075
$dp_{t+1}$	1.342 (2.338)	3.124 (4.218)	0.175 (0.070)	-7.924 (8.378)	-1.040 (4.329)	0.056
$cr_{t+1}$	-0.007 (0.023)	0.045 (0.041)	0.002 (0.001)	0.334 (0.082)	0.054 (0.042)	0.139
$spr_{t+1}$	0.109 (0.047)	0.370 (0.085)	0.000 (0.001)	0.785 (0.168)	0.776 (0.087)	0.340
<b>Restricted parameter estimates</b>						
	$r_{tb,t}$	$r_{nom,t}$	$dp_t$	$cr_t$	$spr_t$	$R^2$
$r_{tb,t+1}$	0.185 (0.043)	1.101 (0.077)	-0.005 (0.001)	0.506 (0.158)	0.409 (0.079)	0.564
$r_{nom,t+1}$				-0.515 (0.190)		0.038
$dp_{t+1}$			0.195 (0.067)			0.031
$cr_{t+1}$			0.002 (0.000)	0.290 (0.065)		0.126
$spr_{t+1}$		0.232 (0.039)		0.635 (0.143)	0.682 (0.042)	0.319

*Notes:* Standard errors in parenthesis. All variables are fractionally differenced using the filter (16). Orders of fractional differentiation correspond to the boldface entries in table 2. The  $R^2$  is based on the fractional differences. The restricted estimates have been obtained from a specification search based on the Schwarz criterium. The unrestricted estimates are obtained by OLS regressions; the restricted estimates are based on FIML.

Table 5: Residual correlations and standard deviations

	$y_{s,t}$	$y_{b,t}$	$r_{tb,t}$	$r_{nom,t}$	$dp_t$	$cr_t$	$spr_t$
$y_{s,t}$	7.706	0.130	0.260	-0.107	-0.963	-0.034	-0.023
$y_{b,t}$		3.812	0.281	-0.641	-0.148	0.707	-0.002
$r_{tb,t}$			0.160	-0.307	-0.294	0.157	0.156
$r_{nom,t}$				0.224	0.113	-0.428	-0.748
$dp_t$					7.910	0.044	0.023
$cr_t$						0.077	-0.056
$spr_t$							0.159

*Notes:* The table reports the standard deviations of the shocks to all series in the system. Residuals for excess bond and stock returns are from the benchmark model in table 3. Residuals for the predictor variables are from the restricted model in table 4. The standard deviations on the diagonal are given as percent per quarter.

Figure 1: Stocks returns and dividends

The figure plots the term structure of risk  $\sigma_y^2(k)$  for excess stock returns using the dividend yield as a predictor. Parameter values are from the left-hand panel in table II on page 240 of Barberis (2000):  $\beta = 0.5118$ ,  $\alpha = 0.9774$ ,  $\sigma_u^2 = 0.0017$ ,  $\sigma_\epsilon^2 = 3.0E-6$ ,  $\sigma_{u\epsilon} = -3.41E-5$ . The monthly results have been annualised in the figure by multiplying with  $\sqrt{12}$ . These parameter values produce the solid line labeled AR1. The "random walk" line is obtained by setting  $\alpha = 1$ . The dashed line uses the fractional model  $(1 - L)^d x_t = \epsilon_t$  with  $d = 0.9$ .

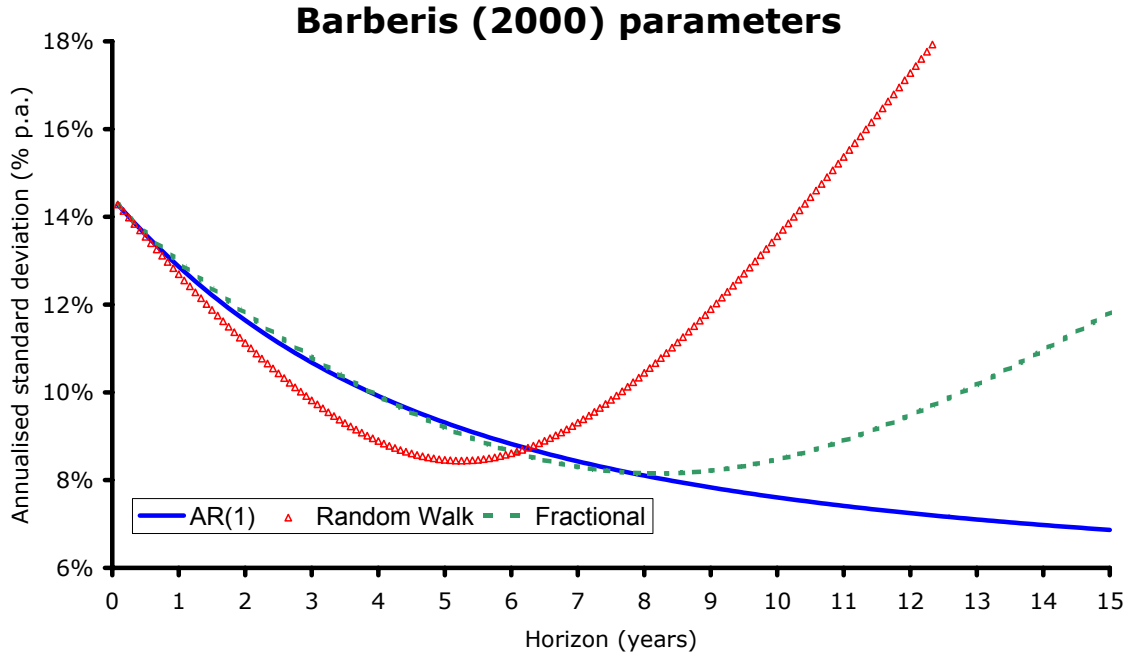


Figure 2: Autocorrelations

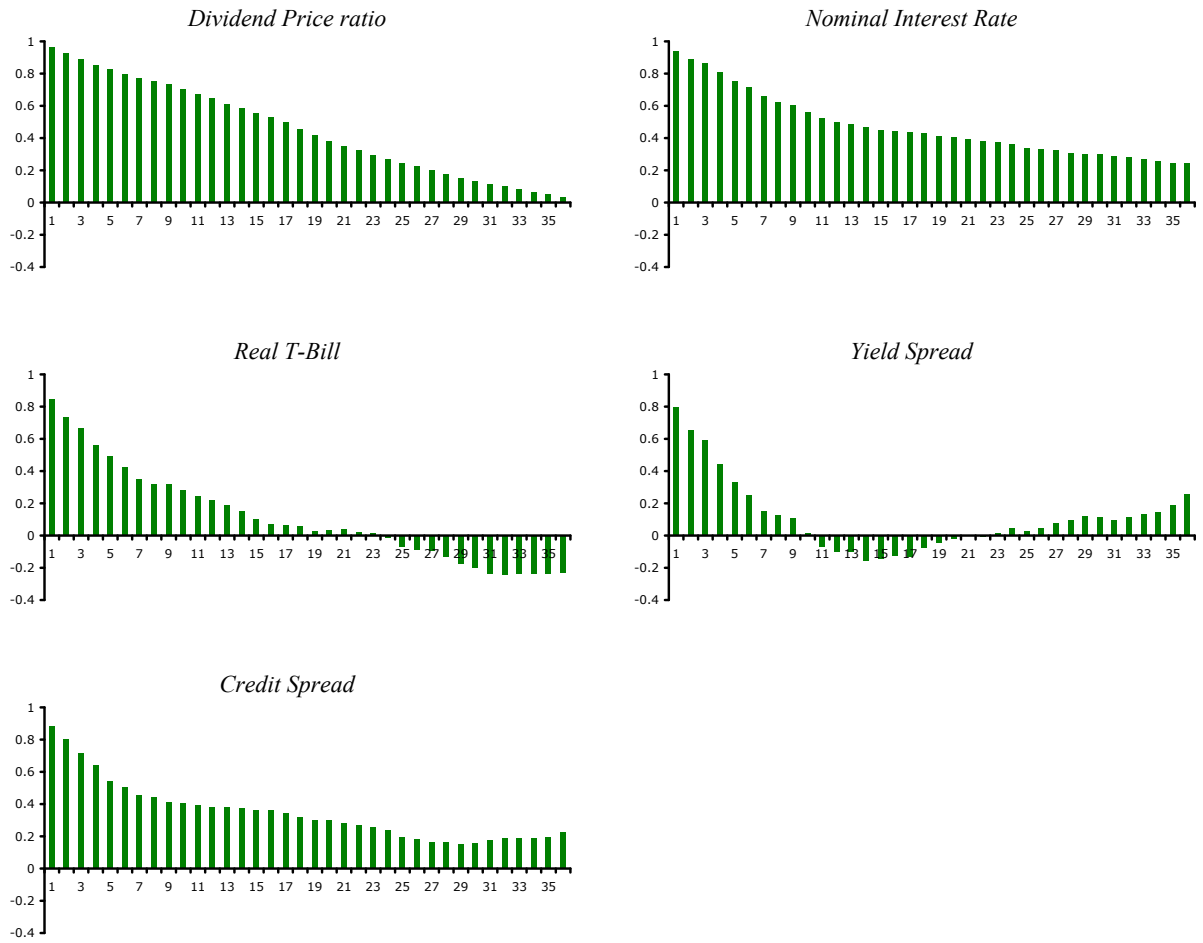




Figure 3: Term structure of risk: benchmark model

The figure plots the term structure of risk for real returns on stocks, bonds and T-Bills. Parameter estimates are based on the prediction estimates labeled *Schwarz* in table 3.

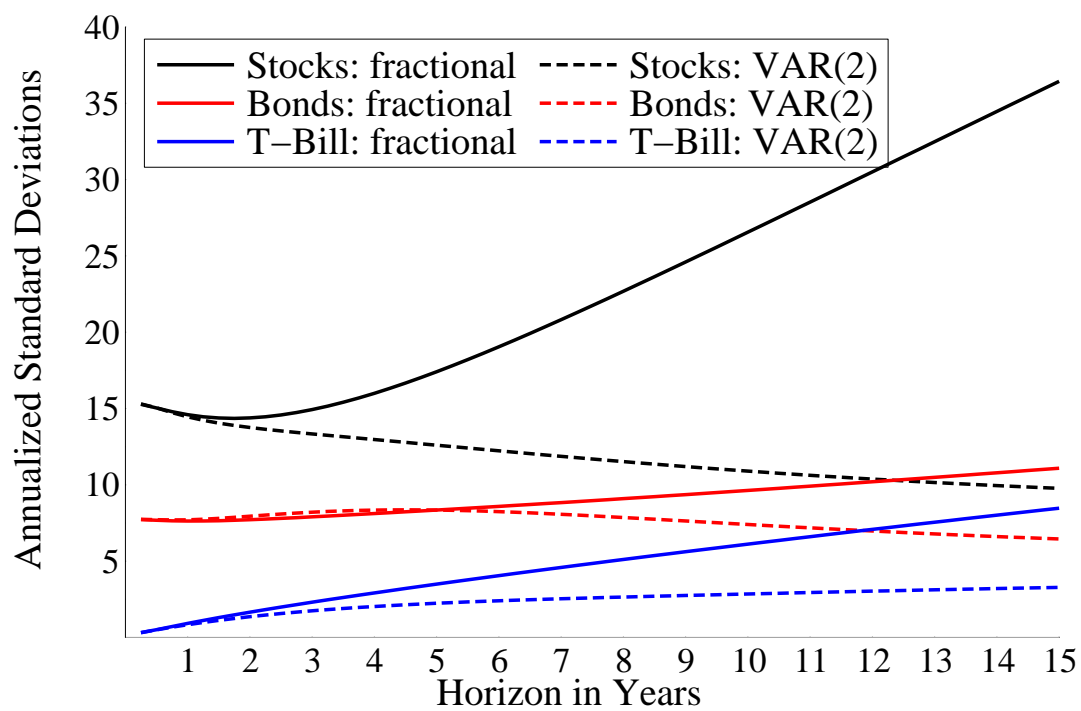
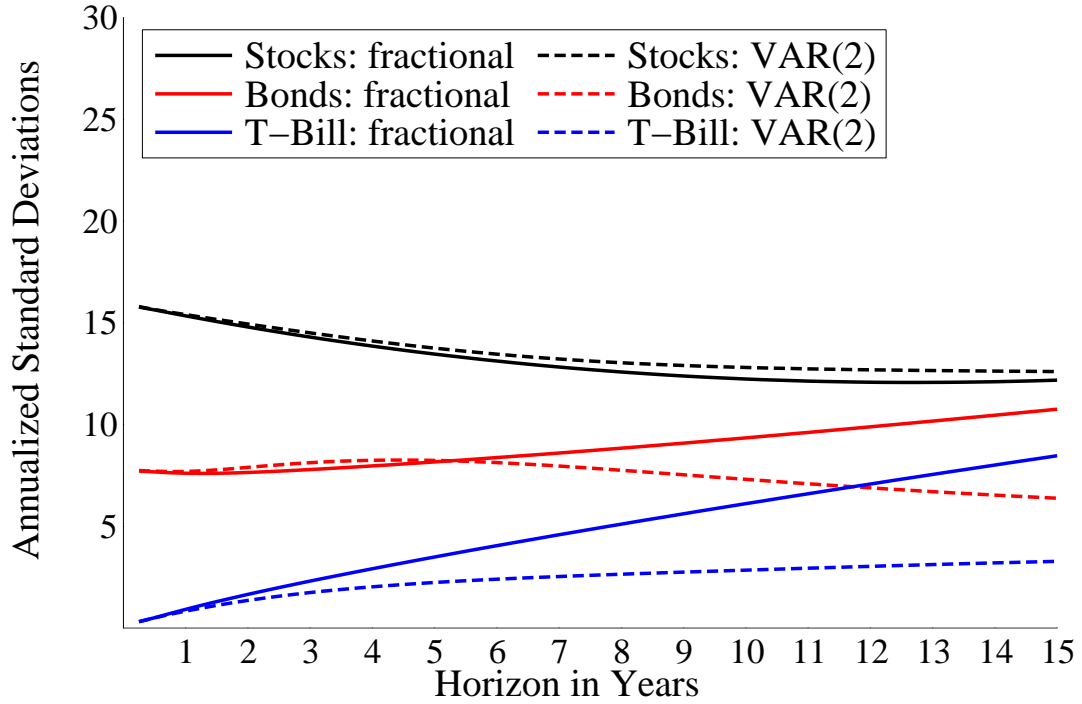


Figure 4: Term structure of risk: restricted models

The figure plots the term structure of risk for real returns on stocks, bonds and T-Bills. Parameter estimates for the top graph are based on the prediction estimates with only the dividend-price ratio (labeled *DivP* in table 3). The bottom graph are the prediction estimates with only the nominal interest rate (*NomR* in table 3).

*Dividend-Price Ratio only*



*Nominal Interest Rate only*

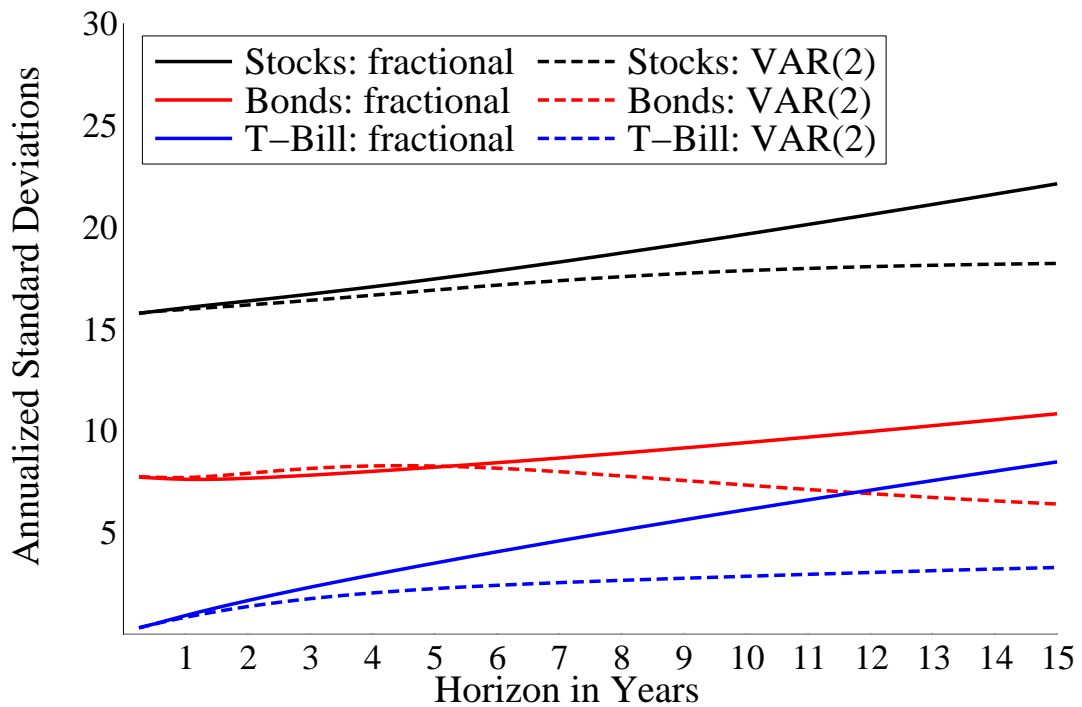
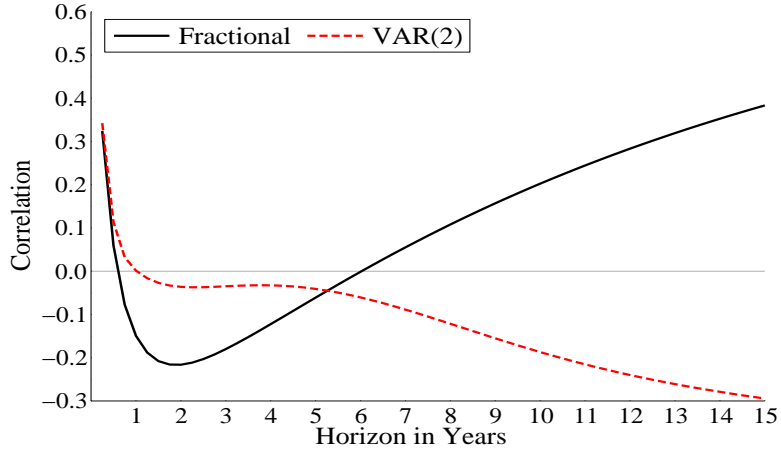
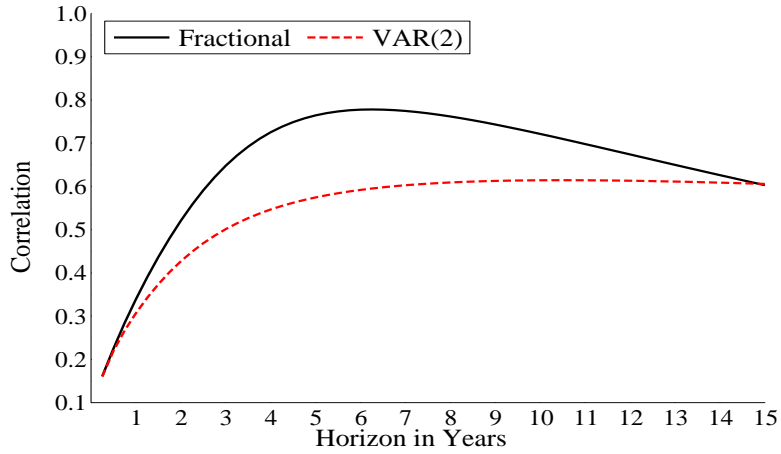


Figure 5: Term structure of correlations

*Bonds and T-Bills*



*Stocks and Bonds*



*Stocks and T-Bills*

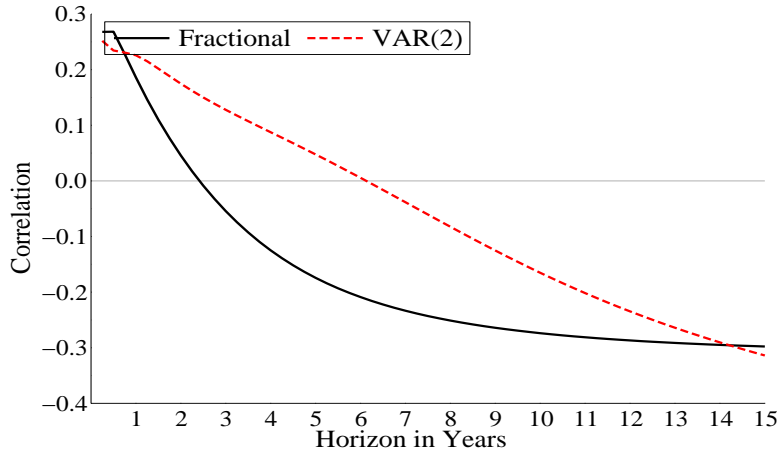


Figure 6: Minimum variance portfolio weights

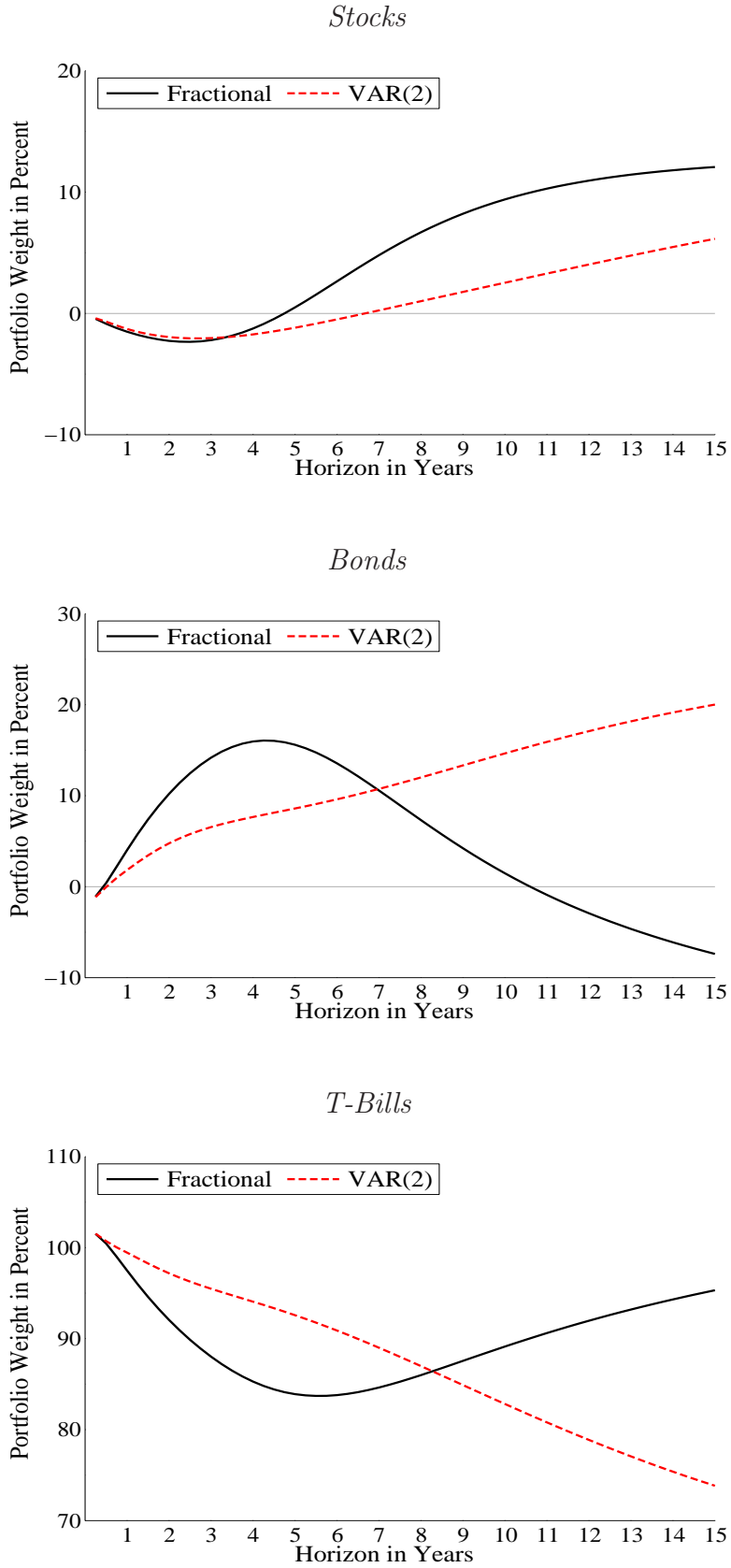


Figure 7: Minimum variance portfolio risk

