Dynamic Bond Portfolios under Model and Estimation Risk

Peter Feldhütter[a] Linda S. Larsen[b] Claus Munk[c] Anders B. Trolle[d]

January 13, 2012

Abstract

We investigate the impact of parameter uncertainty on the performance of bond portfolios. We assume that the data generating process is represented by the well-established three-factor essentially affine Gaussian term structure model. We estimate this model and three simpler models to US data using the Markov Chain Monte Carlo method which provides a posterior distribution of parameters given the data and a point estimate (the median). An investor following the seemingly optimal portfolio strategy for the true model using the parameter point estimate will suffer a utility loss if the true parameters differ from the point estimate, and we find that the average utility loss based on the posterior parameter distribution is big. The degree of parameter uncertainty increases with the number of term structure factors, and we show that investors with moderate or high risk aversion will suffer a smaller average utility loss if they follow the portfolio strategy of a simple one-factor model instead of the more complex true model.

Keywords: Suboptimal investments, parameter uncertainty, utility losses, bond portfolios, MCMC estimation

JEL subject codes: G11

[a] London Business School, Regent's Park, London NW1 4SA, United Kingdom. E-mail: pfeldhutter@london.edu

[b] Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. E-mail: lsn@sam.sdu.dk

[c] Department of Economics and Business & Department of Mathematics, Aarhus University, Bartholins Allé 10, DK-8000 Aarhus C, Denmark. E-mail: cmunk@econ.au.dk

[d] Swiss Finance Institute, Ecole Polytechnique Fédérale de Lausanne, Quartier UNIL-Dorigny, Extranef 216, CH-1015 Lausanne, Switzerland. E-mail: anders.trolle@epfl.ch

We are grateful for comments and suggestions from Christian Riis Flor, Simon Lysbjerg Hansen, Frank de Jong, and Carsten Sørensen.

Dynamic Bond Portfolios under Model and Estimation Risk

The paper contains graphs in color, use color printer for best results.

Abstract

We investigate the impact of parameter uncertainty on the performance of bond portfolios. We assume that the data generating process is represented by the well-established three-factor essentially affine Gaussian term structure model. We estimate this model and three simpler models to US data using the Markov Chain Monte Carlo method which provides a posterior distribution of parameters given the data and a point estimate (the median). An investor following the seemingly optimal portfolio strategy for the true model using the parameter point estimate will suffer a utility loss if the true parameters differ from the point estimate, and we find that the average utility loss based on the posterior parameter distribution is big. The degree of parameter uncertainty increases with the number of term structure factors, and we show that investors with moderate or high risk aversion will suffer a smaller average utility loss if they follow the portfolio strategy of a simple one-factor model instead of the more complex true model.

Keywords: Suboptimal investments, parameter uncertainty, utility losses, bond portfolios, MCMC estimation

JEL subject codes: G11

1 Introduction

A number of recent papers have analyzed dynamic portfolio strategies taking into account stochastic interest rates[1] and time-varying bond risk premia.[2] The consensus is that both features are quantitatively important for investors. For an unconstrained investor, Sangvinatsos and Wachter (2005) find large utility losses associated with portfolio strategies that ignore time-variation in bond risk premia. For a life-cycle investor facing borrowing, short-sales, and liquidity constraints, Koijen, Nijman, and Werker (2010) find lower, but still significant, utility losses from ignoring time-variation in bond risk premia. Other papers find significant utility losses from portfolio strategies that ignore the multi-factor nature of term structure dynamics (see e.g. Larsen and Munk (2012)). Hence, it appears that investors should base their bond portfolio strategies on advanced models of the term structure of interest rates.

The studies cited above ignore estimation risk by assuming that the data generating process is known with certainty, both in terms of model and model parameters. However, model parameters are not known with certainty, and parameter uncertainty increases with the number of term structure factors and with the complexity of the risk premium specification as shown by Duffee (2002). Therefore, even if the investor has identified the true data generating model he will suffer a utility loss if he applies the seemingly optimal portfolio strategy using his parameter estimate instead of the true, unknown parameter set. Moreover, it is possible that by basing his portfolio choice on a more parsimonious – but misspecified – model with less parameter uncertainty, the investor suffer a smaller utility loss. It is this trade-off between advanced models subject to a significant amount of parameter uncertainty and more robust models subject to less parameter uncertainty that we analyze in this paper.

More specifically, we assume that the true data generating model is well approximated by a three-factor Gaussian term structure model where bond risk premia are affine in the factors. This model has been shown by Dai and Singleton (2002) and Duffee (2002) to outperform other models in terms of capturing the predictability in bond returns. The same data generating model is assumed by Sangvinatsos and Wachter (2005), whereas

[1] See, e.g., Sørensen (1999), Brennan and Xia (2000), Campbell and Viceira (2001), Wachter (2003), and Munk and Sørensen (2004).

[2] See, e.g., Sangvinatsos and Wachter (2005) and Koijen, Nijman, and Werker (2010).

Koijen, Nijman, and Werker (2010) assume a two-factor version of the model. We then consider reducing the complexity of the model along two dimensions: we reduce the number of factors from three to one and we reduce the risk premium specification from being affine in the factors to being constant. This leads to three increasingly parsimonious models: a three-factor model with constant bond risk premia and two one-factor models in which bond risk premia are affine and constant, respectively.

For concreteness and tractability, we follow the standard in the portfolio choice literature by assuming that the investor has CRRA utility over terminal wealth. In each of the four term structure models we obtain semi-analytical expressions for the dynamic portfolio strategy which is optimal in the absence of model and estimation risk. We assume that the investor ignores parameter and model uncertainty and follows the portfolio strategy associated with the model which the investors thinks is correct, and the investor does so using his point estimates of the parameters of that model. We make this assumption for two reasons. First, it is the typical assumption in the dynamic portfolio choice literature, and we want to investigate the extent to which existing results hold up in the presence of parameter uncertainty. Second, in a dynamic setting it is very complicated to incorporate parameter and model uncertainty into the decision making, except in very stylized settings.[3]

All four models are estimated by Markov Chain Monte Carlo (MCMC) on a panel data set of U.S. Treasury yields with daily observations over the period 1971-2006.[4] MCMC is convenient as it provides the posterior distribution of model parameters given the data. This is used for evaluating the robustness of the portfolio strategies implied by the different models and for quantifying the utility losses due to parameter uncertainty.

Consistent with virtually all empirical studies of dynamic term structure models (see, e.g. Dai and Singleton (2002) and Duffee (2002)), we find that market price of risk parameters are imprecisely estimated; confidence intervals are wide and often parameters are not statistically significant. As expected, we find the widest confidence intervals in case of the

[3]A few papers studies dynamic portfolio choice with parameter uncertainty (see, e.g., Brennan (1998), Barberis (2000), and Xia (2001)) and model uncertainty (see, e.g., Maenhout (2004, 2006)) but the tractability in those papers is predicated upon assuming very simple data-generating models.

[4]We are not the first to use MCMC to estimate term structure models, see e.g. Sarno, Schneider, and Wagner (2011), Feldhütter (2008), Kaminska, Vayanos, and Zinna (2011), and Ang, Dong, and Piazzesi (2007).

most complex models. Since the portfolio allocations are sensitive to market price of risk parameters, it follows that the allocations are associated with wide confidence intervals as well. Often it is not even clear whether the investor should take a short or a long position in a given bond. Again, allocations for the most complex models are associated with the widest confidence intervals. This suggests that it may be profitable to follow the more robust portfolio strategies implied by more parsimonious models.

To investigate this issue we take a Bayesian approach. For a suboptimal portfolio strategy, we can compute the utility loss analytically, conditional on knowing the parameters of the data generating model (the three-factor model with affine bond risk premia). However, the parameters of this model are uncertain. To obtain the expected utility loss taking parameter uncertainty into account, we integrate the loss over the posterior distribution of the parameters. A similar approach is used in a very different setting by Korteweg and Polson (2009) who analyze the effect of parameter uncertainty on corporate credit spreads.

We find that parameter uncertainty leads to significant utility losses and that these losses increase with the complexity of the model. In fact, long-term investors with moderate to high risk aversion are often better off basing their portfolio decisions on more parsimonious models, since utility losses for these models due to both parameter uncertainty and model misspecification tend to be lower than utility losses for the true data generating model solely due to parameter uncertainty. For instance, assuming a relative risk aversion of 5 and an investment horizon of 5 years – rather typical values – and abstracting from parameter uncertainty, an investor using a three-factor (one-factor) model with constant risk premia will suffer a utility loss of 50 (49) percent – consistent with findings in the papers cited above that suboptimal investment strategies carry large costs. However, this conclusion no longer holds true once parameter uncertainty is taken into account. In this case an investor using the correct three-factor model with affine risk premia will suffer an average utility loss of 61 percent, whereas an investor using a three-factor (one-factor) model with constant risk premia will suffer an average utility loss of 59 (51) percent. Hence, the suboptimal investment strategy based on the one-factor model with constant risk premia in fact carries a 20% lower utility loss compared to the investment strategy based on the true model. For a higher level of risk aversion or a longer investment horizon the difference in the average utility loss becomes even bigger. For example, for a risk aversion of 10 and an investment horizon of 5 years an investor using the correct model will suffer an average utility loss of 66%, whereas an investor basing his investment strategy on the most parsimonious – and misspecified – model only suffers a loss of 34%, a loss reduction of approximately 48%.

Our paper differs from other papers in the literature investigating the impact of model and parameter uncertainty on portfolio strategies. To the best of our knowledge, our paper is the first to analyze the quantitative effect of parameter uncertainty in portfolio strategies, i.e. how much do the investor actually suffers due to estimation risk. Furthermore, this paper investigates the impact of parameter uncertainty in portfolios of bonds, whereas earlier papers mainly focus on stock portfolios. Klein and Bawa (1976) is one of the first papers to study the effect of estimation risk on optimal portfolio strategies. They incorporate estimation risk directly into the decision process of the investor and determine its effects on optimal portfolio choice under uncertainty. Barry and Brown (1985) propose a simple model of equilibrium asset pricing in a setting with parameter uncertainty. Barberis (2000) incorporate parameter uncertainty in a setting with predictability in asset returns. He shows, in line with other papers incorporating model and parameter uncertainty, that investors should take less extreme positions in the risky assets if the model parameters are associated with uncertainty.[5] All of these papers take the estimation risk into account in their portfolio decision, whereas our paper look at the consequences of an investor who ignores parameter uncertainty and simply follows the optimal investment strategy determined by theory.

The remainder of the paper is organized in the following way. Section 2 sets up the modeling framework, and specifies the investment strategies. Section 3 discusses the data and the estimation procedure. Section 4 contains the results. Section 5 summarizes and concludes.

2 The general setup

We consider an arbitrage-free economy over the time interval $[0, T]$ where trading takes place continuously in time. We assume that the true data generating process is well

[5] See also Brennan (1998), Xia (2001), Maenhout (2004, 2006), and Johannes, Korteweg, and Polson (2011).

approximated by a three-factor Gaussian term structure model where bond risk premia are affine in the factors. This model has been shown by Dai and Singleton (2002) and Duffee (2002) to outperform other models in terms of capturing the predictability in bond returns.

We want to analyze the effect of reducing the complexity of the model along two dimensions. That is, reducing the number of factors from three to one, and reducing the risk premium specification from being affine in the factors to be being constant. This leads to three increasingly parsimonious models that also need to be specified: a three-factor model with constant bond risk premia, and two one-factor models where bond risk premia are affine and constant, respectively. The true model as well as the three increasingly parsimoniously models are all nested by the following m-factor Gaussian term structure model where bond risk premia are affine in the factors. Let r_t denote the instantaneous real risk-free rate and assume that

$$r_t = \delta_0 + \delta_X' X_t, \tag{1}$$

where δ_0 is a constant, δ_X is an $m \times 1$ vector, and $X_t = (X_{1t}, X_{2t}, \ldots, X_{mt})'$ is an $m \times 1$ vector of state variables that follows the process

$$dX_t = \kappa\left(\theta - X_t\right) dt + \sigma_X \, dz_t \tag{2}$$

under the physical measure $\mathbb{P}$, where $z = (z_t)_{t\in[0,T]}$ is a standard m-dimensional standard Brownian motion. The $m \times m$ constant matrix σ_X is assumed invertible and determines the variance-covariance matrix of the state variables over the next instant, $\sigma_X \sigma_X'$, κ is an invertible $m \times m$ matrix, and θ is an $m \times 1$ vector. Furthermore, the market price of risk associated with the shock process z is assumed to be linear in X, i.e.

$$\lambda_t = \lambda_0 + \lambda_X X_t, \tag{3}$$

where λ_0 is an $m \times 1$ vector and λ_X is an invertible $m \times m$ matrix. For the models with a constant bond risk premia λ_X is set equal to zero. For later use, we need the dynamics of the state variables under the risk-neutral probability measure $\mathbb{Q}$

$$dX_t = \tilde{\kappa}\left(\tilde{\theta} - X_t\right) dt + \sigma_X \, dz_t^{\mathbb{Q}}, \tag{4}$$

where $z = (z_t^{\mathbb{Q}})_{t\in[0,T]}$ is a standard Brownian motion under $\mathbb{Q}$ with $dz_t^{\mathbb{Q}} = dz_t + \lambda_t dt$. Furthermore, $\tilde{\kappa} = \kappa + \sigma_X \lambda_X$ and $\tilde{\theta} = \tilde{\kappa}^{-1}\left(\kappa\theta - \sigma_X \lambda_0\right)$.

The investor can invest in an instantaneously risk-free asset, interpreted as short-term cash deposits, which yields the continuously compounded rate of return r_t. Besides the risk-free asset the investor can invest in $n < \infty$ zero-coupon bonds. As shown by Duffie and Kan (1996) the price of a zero-coupon bond maturing at T takes the form $P_t^T = P^T(t, X_t)$ where

$$P^T(t, X) = \exp\left\{-A(T-t) - B(T-t)'X\right\},$$

and $A : [0, T] \to \mathbb{R}$, $B : [0, T] \to \mathbb{R}^m$ are solutions to the system of ordinary differential equations:

$$\frac{\partial B(\tau)}{\partial \tau} = \delta_X - \tilde{\kappa}' B(\tau) \tag{5}$$

$$\frac{\partial A(\tau)}{\partial \tau} = B(\tau)' \tilde{\kappa}\, \tilde{\theta} - \frac{1}{2} B(\tau)'\, \sigma_X\, \sigma_X'\, B(\tau)' + \delta_0, \tag{6}$$

These equations can easily be solved with the boundary conditions $A(0) = B(0) = 0$. The dynamics of the zero-coupon bond price with maturity T follows from Ito's lemma:

$$\frac{dP_t^T}{P_t^T} = \big(r_t - B(T-t)'\sigma_X \lambda_t\big)\, dt - B(T-t)'\sigma_X dz_t. \tag{7}$$

2.1 The investor

The investor chooses an investment strategy, which we represent by the n-dimensional continuous-time process $\pi = (\pi_t)$, where $\pi_t = (\pi_{1t}, \pi_{2t}, \ldots, \pi_{nt})'$ is the vector of fractions of wealth ("portfolio weights") invested in the different zero-coupon bonds at time t. The remaining fraction of wealth $1 - \pi_t' \mathbf{1}$ is invested in the instantaneously risk-free asset. We ignore intermediate consumption and income other than financial returns. Hence given a positive initial wealth W and an investment strategy π the investor's wealth will satisfy the self-financing condition

$$dW_t = W_t \left[r_t - \pi_t' \mathcal{B}(\tilde{\tau})' \sigma_X \lambda_t\right] dt - W_t \pi_t' \mathcal{B}(\tilde{\tau})' \sigma_X\, dz_t, \tag{8}$$

where $\mathcal{B}(\tilde{\tau})$ is an $m \times n$ matrix with the i'th column representing the B-vector associated with the i'th zero-coupon bond, i.e.

$$\mathcal{B}(\tilde{\tau}) = \left(B(T_1 - t), B(T_2 - t), \ldots, B(T_n - t)\right).$$

We assume that the investor is concerned with maximizing expected utility of wealth at some future date T and that the utility function is of the CRRA type. We further assume that the investor acts as if he knows model parameters and the state vector. Then, the indirect utility is given as

$$J(W,X,t) = \sup_{(\pi_s)_{s\in[t,T]}} \begin{cases} \mathbb{E}_{W,X,t}\left[\frac{1}{1-\gamma} W_T^{1-\gamma}\right], & \gamma > 1, \\ \mathbb{E}_{W,X,t}\left[\ln W_T\right], & \gamma = 1, \end{cases} \tag{9}$$

where $\mathbb{E}_{W,X,t}$ denotes the expectation operator given $W_t = W$ and $X_t = X$ under the physical measure $\mathbb{P}$, and γ is the constant relative risk aversion parameter.[6] The optimal investment strategy π^* is the one satisfying (9). It is well known that the optimal investment strategy for a CRRA investor will be independent of his wealth level. Hence, we will focus on strategies of the form $\pi_t = \pi(X_t, t)$.

2.2 The optimal investment strategy

Sangvinatsos and Wachter (2005) derive a semi-analytical expression for the optimal portfolio strategy and the associated expected utility.[7] This involves trading in $n = m$ bonds. For completeness, we state the result here:

Proposition 1 *The expected utility generated by the optimal investment strategy π^* is*

$$J(W,X,t) = \begin{cases} \frac{1}{1-\gamma}\left(W\, e^{F(T-t)+H(T-t)'X+\frac{1}{2}X'G(T-t)X}\right)^{1-\gamma}, & \gamma > 1, \\ \ln W + F(T-t) + H(T-t)'X + \frac{1}{2}X'G(T-t)X, & \gamma = 1, \end{cases} \tag{10}$$

where $F(\tau)$, $H(\tau)$, and $G(\tau)$ solve a system of ordinary differential equations given in Appendix A. The optimal strategy is

$$\begin{aligned} \pi^*(X,t) = &-\frac{1}{\gamma}\left(\sigma_X'\mathcal{B}(\tilde{\tau})\right)^{-1}(\lambda_0 + \lambda_X X) \\ &+ \frac{\gamma-1}{\gamma}\left(\sigma_X'\mathcal{B}(\tilde{\tau})\right)^{-1}\sigma_X'\left(H(\tau) + \frac{1}{2}\left(G(\tau)+G(\tau)'\right)X\right). \end{aligned} \tag{11}$$

The optimal investment strategy is composed of two portfolios: a speculative portfolio (the first term in (11)) and a hedge portfolio (the second term in (11)). The hedge portfolio describes how the investor should optimally hedge against the changes in the investment opportunity set as a result of stochastic variation in the short rate and the market prices

[6] We assume $\gamma \geq 1$ to avoid problems with infinite expected utility that may arise for $0 < \gamma < 1$, cf. Kim and Omberg (1996) and Korn and Kraft (2004).

[7] Sangvinatsos and Wachter (2005) also consider inflation uncertainty, a dimension that we abstract from here.

of risk. The hedge portfolio consists of two components – the first is due to the stochastic variation in the short rate, while the second is due to the stochastic variation in the market price of risk vector. If market prices of risk are constant – that is, $\lambda_X = 0$ – then $G = 0$ and the second term in the hedge portfolio disappear. A further discussion of the investment strategy is given in Sangvinatsos and Wachter (2005).

2.3 Suboptimal investment strategies

The true data generating model is assumed to be well approximated by the three-factor essentially affine model stated above with $m = 3$. However, the model parameters describing the model are not known with certainty. Hence, even if the investor has identified the true data generating model he might follow a suboptimal investment strategy, due to the fact that he bases his strategy on the wrong set of parameters. It might even be the case that an investment strategy based on a more parsimonious – but misspecified – model with less parameter uncertainty, outperforms the strategy based on the true model due to parameter uncertainty.

It is this trade-off between basing the investment strategy on advanced models subject to a significant amount of parameter uncertainty or more robust models subject to less parameter uncertainty that we want to analyze. In particular, we will consider the following four investment strategies: (i) an investment strategy based on the true three-factor essentially affine model with affine bond risk premia, (ii) an investment strategy based on a three-factor completely affine model with constant bond risk premia, (iii) an investment strategy based on a one-factor essentially affine model with affine bond risk premia, and (iv) an investment strategy based on a one-factor completely affine model with constant bond risk premia. All four investment strategies can be nested by the investment strategy stated in Proposition 1, that is

$$\begin{aligned}\pi(X,t) = &-\frac{1}{\gamma}\left(\hat{\sigma}_X'\hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\left(\hat{\lambda}_0 + \hat{\lambda}_X X\right) \\ &- \frac{1-\gamma}{\gamma}\left(\hat{\sigma}_X'\hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{\sigma}_X'\left(\hat{H}(\tau) + \frac{1}{2}\left(\hat{G}(\tau) + \hat{G}(\tau)'\right)X\right).\end{aligned} \tag{12}$$

It is very important to clarify that each of the four investment strategies are based on a given set of parameter estimates. In particular, the hat ($\hat{}$) on the parameters indicates that not the true set of parameters is used. Furthermore, the functions $\hat{G}$ and $\hat{H}$ which solve the system of ODEs given in Proposition 1 is determined under the assumption of one of the four investment strategies and the given set of parameter estimates. We can evaluate these strategies according to the following proposition. A sketch of the proof can be found in Appendix A.

Proposition 2 *Assuming model* (1)-(3) *the expected utility generated by the investment strategy stated in* (12) *is given by*

$$\hat{J}(W,X,t) = \begin{cases} \frac{1}{1-\gamma}\left(W\, e^{C_1(T-t)+C_2(T-t)'X+\frac{1}{2}X'C_3(T-t)X}\right)^{1-\gamma}, & \gamma > 1, \\ \ln W + C_1(T-t) + C_2(T-t)'X + \frac{1}{2}X'C_3(T-t)X, & \gamma = 1, \end{cases} \tag{13}$$

where the deterministic functions $C_1(\tau)$*,* $C_2(\tau)$*, and* $C_3(\tau)$ *solve a system of ordinary differential equations given in Appendix A.2.*

By definition, with the same initial wealth, a suboptimal investment strategy will generate a lower level of expected utility than the optimal investment strategy. We define the loss from following the suboptimal strategy as the fraction of initial wealth that the investor would be willing to give up to be able to invest according to the optimal strategy instead of the suboptimal strategy. We refer to this loss as the utility loss associated with the suboptimal investment strategy and denote it by L. By definition it solves

$$\hat{J}(W,X,t) = J(W(1-L),X,t), \tag{14}$$

and straightforward calculations using (10) and (28) show that

$$L \equiv L(X,\tau) = 1 - e^{C_1(\tau)-F(\tau)+(C_2(\tau)-H(\tau))'X+\frac{1}{2}X'(C_3(\tau)-G(\tau))X} \tag{15}$$

for $\gamma > 1$. The same formula holds for $\gamma = 1$, i.e. (15) holds for all values of $\gamma \geq 1$.

3 Estimation

To compare the four different suboptimal investment strategies we need to estimate all four models. First, we need to estimate the true data generating model, that is a three-factor Gaussian term structure model in which bond risk premia are affine in the factors. This model has been shown by Dai and Singleton (2002) and Duffee (2002) to outperform other models in terms of capturing the predictability in bond returns. Second, we need to estimate the three increasingly parsimonious models: the three-factor Gaussian model with constant bond risk premia, and the two one-factor models in which bond risk premia are affine and constant, respectively. To make a fair comparison of the models we estimate all four models by the same estimation method and on the same data set.

3.1 Estimation procedure

To avoid overidentification we apply the parametrization of Dai and Singleton (2000) and assume that

$$\tilde{\theta} = \tilde{\kappa}^{-1}\left(\kappa\theta - \sigma_X\lambda_0\right) = 0,$$

where σ_X equals the $(m \times m)$-identity matrix, and $\tilde{\kappa} = \kappa + \sigma_X\lambda_X$ is a $(m \times m)$-lower triangular matrix. For the completely affine models we put $\lambda_X = 0$. We adopt a Bayesian approach and estimate the models by Markov Chain Monte Carlo (MCMC) as proposed by Eraker (2001)[8]. The use of MCMC is chosen for two reasons. First, we are able to simultaneously estimate parameters and latent variables given the observed data. Second, MCMC allows us to quantify the uncertainty present in estimating parameters or state variables. An investor's optimal portfolio choice can be very sensitive to the parameters being used, and the use of MCMC enables us to quantify this risk.

At time $t = 1, ..., T$ we observe k yields which are stacked in a k-vector

$$Y_t = (Y(t, \tau_1), ..., Y(t, \tau_k))'.$$

The yields are all observed with a measurement error

$$Y_t = A + BX_t + \epsilon_t$$

where A is a k-vector and B is a $k \times m$ matrix. We assume that the measurement errors are independent and normally distributed with zero mean and common variance such that

$$\epsilon_t \sim N(0, D), \qquad D = \varphi^2 I_k,$$

where I_k denote the $k \times k$ identity matrix. To simplify the notation in the following, we denote

$$\Theta^Q = \left(\tilde{\kappa}, \delta_0, \delta_X\right), \qquad \Theta^P = \left(\lambda_0, \lambda_X\right),$$

[8] For a general introduction to MCMC, see Robert and Casella (2004) and for a survey of MCMC methods in financial econometrics see Johannes and Polson (2006).

and $\Theta = (\Theta^Q, \Theta^P, D)$.

We are interested in samples from the target distribution $p(\Theta, X|Y)$. The Hammersley-Clifford Theorem (Hammersley and Clifford, 1970, and Besag, 1974) implies that samples are obtained from the target distribution by sampling from a number of conditional distributions. Effectively, MCMC solves the problem of simulating from a complicated target distribution by simulating from simpler conditional distributions. If one samples directly from a full conditional distribution, the resulting algorithm is the Gibbs sampler (Geman and Geman, 1984). If it is not possible to sample directly from the full conditional distribution, one can sample by using the Metropolis-Hastings algorithm (Metropolis et al., 1953). We use a hybrid MCMC algorithm that combines the Gibbs sampler and the Metropolis-Hastings algorithm since not all the conditional distributions are known. Specifically, the MCMC algorithm is given by[9]

$$\begin{aligned} p(\Theta^Q|\Theta^P, D, X, Y) &\sim \text{Metropolis-Hastings} \\ p(\lambda_0|\Theta_{\setminus\lambda_0}, X, Y) &\sim \text{Normal} \\ p(\lambda_X|\Theta_{\setminus\lambda_X}, X, Y) &\sim \text{Normal} \\ p(D|\Theta_{\setminus D}, X, Y) &\sim \text{Inverse Wishart} \\ p(X|\Theta, Y) &\sim \text{Metropolis-Hastings} \end{aligned}$$

Details in the derivations of the conditionals and proposal distributions in the Metropolis-Hastings steps are given in Appendix B.1. Both the parameters and the latent processes are subject to constraints, and if a draw is violating a constraint it can simply be discarded (Gelfand et al., 1992).

3.2 Data

We use daily (continuously compounded) 1, 2, 3, 5, 7, and 10-year zero-coupon yields extracted from prices on off-the-run US Treasury securities for the period from August 16, 1971 to August 21, 2006.[10] Figure 1 displays the data.

[9]Here $\Theta_{\setminus a}$ denotes the parameter vector excluding the parameter a.

[10]Off-the-run securities are defined as securities that are not among the two most recently issued securities with maturities of two, three, four, five, seven, and ten years. The data set is discussed in detail in Gürkaynak et al. (2006) and is posted on the website http://www.federalreserve.gov/pubs/feds/2006.

[Figure 1 about here.]

3.3 Parameter estimates

In estimating each model we use an algorithm calibration period of 3 million draws, a burn-in period of 5 million draws and an estimation period of 5 million draws. We keep every 5,000'th draw in the estimation period, which leaves 1,000 draws. For each of the four models we find our benchmark estimates among the 1,000 draws as follows. Following Collin-Dufresne et al. (2008), let ϕ_i denote the ith parameter draw and let $\tilde{\phi}_i$ denote the same vector normalized by the posterior standard deviations. The benchmark estimate is the draw i minimizing:

$$\sum_j \left| \tilde{\phi}_j - \tilde{\phi}_i \right| .$$

This version of the multivariate posterior median ensures that parameter restrictions are satisfied for our parameter estimates, which might not be the case if the point estimates are based on univariate medians. For each parameter, we report the benchmark estimate along with a univariate confidence band based on the 2.5% and the 97.5% percentile of the 1,000 MCMC draws of the posterior distribution.

Tables 1, 2, and 3 display parameter estimates along with their confidence intervals for the four models considered in the paper. It is interesting to note that the market price of risk parameters are generally very imprecisely estimated; the confidence intervals are wide and many of the parameters are not statistically significant. The next section explores the implications for portfolio choice.

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

4 Results

The MCMC estimation technique allows us to quantify the uncertainty inherent in estimating parameters and state variables. This implies that we can also quantify the uncertainty in the portfolio weights induced by the uncertainty in the parameter estimates. Given the empirical evidence, the three-factor essentially affine model is the model we refer to as our true model. Hence we assume that we know the true model, but we do not know the true set of parameter estimates. Each draw from the posterior distribution of the three-factor essentially affine model is a guess of the true parameters.

We consider four different types of investors. The first investor knows the true model and hence bases his investment strategy on that. However, he does not know the true set of parameter estimates but uses the best set of estimates he has – that is the posterior median parameters given in Table 1. The second investor has the number of factors describing the economy right, but misses the fact that the bond risk premia are linear in the state variable. That is, he bases his investment strategy on the three-factor completely affine model and uses the posterior median parameters given in Table 2. The third investor knows that the bond risk premia are linear in the state variables, but misses the right number of factors describing the economy. That is, he bases his investment strategy on the one-factor essentially affine model and uses the posterior median parameters given in Panel A in Table 3. Finally, the fourth investor both misses the number of factors describing the economy as well as the linearity of bond risk premia. That is, he bases his investment strategy on the one-factor completely affine model with the parameter estimates given in Panel B in Table 3.

We assume that the investor at every point in time has access to trade in a 1-year, 5-year, and 10-year zero-coupon bond as well as the risk-free asset. The investors who believe that one factor only is needed to describe the economy will trade in one of the three bonds. If the economy truly was described by a single state variable, the expected utility from following the investment strategy would be independent of the time-to-maturity of the bond. However, in this setting the true model of the world is a three-factor model. In a different setup Brennan and Xia (2002) show that if an investor is allowed to trade only one bond, the optimal maturity of that bond depends on the investor's investment horizon. Hence the loss an investor suffers from following a one-factor strategy may depend on the time-to-maturity of the bond. We will investigate this later on but for now, we assume that the investors who follow a one-factor strategy invests in a 5-year-to-maturity bond as well as the risk-free asset.

4.1 Investment strategies

The investor who bases his investments on the true model, the essentially affine three-factor model, follows the strategy given in (11) with $m = 3$, i.e.

$$\pi(X,t) = -\frac{1}{\gamma}\left(\hat{\sigma}'_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\left(\hat{\lambda}_0 + \hat{\lambda}_X X\right) - \frac{1-\gamma}{\gamma}\left(\hat{\sigma}'_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{\sigma}'_X\left(\hat{H}(\tau) + \frac{1}{2}\left(\hat{G}(\tau) + \hat{G}(\tau)'\right)X\right). \tag{16}$$

The investor uses the benchmark estimates given in Table 1. Table 4 displays the portfolio weights for four different investment horizons and the associated 95% confidence intervals calculated by the use of each of the draws from the posterior distribution of the parameter estimates. Panel A shows the portfolio weights for an investor with a risk aversion of $\gamma = 5$, while Panel B shows the weight for an investor with a risk aversion of $\gamma = 10$. The highly leveraged portfolio weights are due to the high correlations among the three bonds. Similar results are shown in Sangvinatsos and Wachter (2005). In both panels we see quite extreme confidence intervals. A big part of the weights are not even statistically significant. In both panels we see that the longer the investment horizon is, the wider the intervals will be. The size of the confidence interval can be seen as the sum of the confidence interval for the speculative portfolio and the hedge portfolio. The confidence interval for the speculative portfolio equals the interval we have for $T = 0$. For $T > 0$ the investor also invests in the hedge portfolio, which implies that the size of the confidence interval gets wider. The same story goes for the investor's risk aversion. We see that the confidence intervals are slightly narrower for the more risk averse investor. A more risk averse investor takes a lower position in the speculative portfolio, which reduces the size of the confidence interval. Even though the size of the confidence intervals is slightly narrower for $\gamma = 10$, we still see intervals running from -11.4 to 1.4 in Panel B.

[Table 4 about here.]

The three-factor completely affine model is given in (1)-(3) with $\lambda_X = 0$ and $m = 3$. An investor basing his investment strategy on this model will follow the strategy in (11) with $G(\tau) = \lambda_X = 0$, i.e.

$$\pi(X,t) = -\frac{1}{\gamma}\left(\hat{\sigma}'_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{\lambda}_0 - \frac{1-\gamma}{\gamma}\left(\hat{\sigma}'_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{\sigma}'_X \hat{H}(\tau), \tag{17}$$

The investor uses the parameter estimates given in Table 2. Table 5 shows the optimal portfolio weight for different horizons and the associated 95% confidence intervals calculated by the use of each of the draws from the posterior distribution of the parameter estimates of the three-factor completely affine model. Panel A shows the portfolio weights for an investor with a risk aversion of $\gamma = 5$, while Panel B shows the weights for an investor with a risk aversion of $\gamma = 10$. Again we get highly leveraged portfolio weights due to the high correlations among the three bonds. Comparing the portfolio weights and confidence intervals with the corresponding optimal portfolio weights and confidence intervals in Table 4 we see the following: The difference in the portfolio weights for a myopic investor (similar to an investor with an investment horizon of $T = 0$) is quite small, while the difference increases for longer investment horizons. This is in line with the findings Sangvinatsos and Wachter (2005), who report that an investor in an essentially affine world has larger hedging demands compared to an investor in a completely affine world. We still see quite wide confidence intervals. However the width is reduced compared to the corresponding confidence intervals in the optimal setup, in particular for longer investment horizons. The reduction is due to the fact that in the completely affine models we have $\lambda_X = 0$. Hence, the investment strategy depends on fewer parameters compared to the essentially affine setup, which reduces the size of the confidence intervals. Furthermore, an investor following the true essentially affine model has larger hedging demands compared to an investor following the completely affine model. This reduces the size of the confidence intervals in the completely affine model even further.

[Table 5 about here.]

The one-factor essentially affine model is given in (1)-(3) with $m = 1$. An investor basing his investment strategy on this model will follow the investment strategy

$$\pi(r,t) = -\frac{1}{\gamma}\left(\hat{\mathcal{B}}(\tilde{\tau})\hat{\sigma}_X\hat{\delta}_X\right)^{-1}\hat{\lambda} - \frac{1-\gamma}{\gamma}\left(\hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\left(\hat{H}(\tau) + \hat{G}(\tau)r\right), \tag{18}$$

where $\hat{\lambda} = \left(\hat{\lambda}_0 - \frac{\hat{\lambda}_X}{\hat{\delta}_X}\hat{\delta}_0\right) + \frac{\hat{\lambda}_X}{\hat{\delta}_X} r$.[11] The investor bases his investment strategy on the parameter estimates given in Panel A in Table 3. The investor who believes that the one-factor essentially affine model gives a true picture of the economy will take the positions given in Table 6. The table displays the optimal portfolio weight for different horizons and the associated 95% confidence intervals calculated by the use of each of the draws from the posterior distribution of the parameter estimates for the one-factor essentially affine model. Panel A shows the portfolio weights for an investor with a risk aversion of $\gamma = 5$, while Panel B shows the weights for an investor with a risk aversion of $\gamma = 10$. Due to the fact that the investor believes that the state of the economy is fully described by a single factor, he only invests in a single bond, here assumed to be a 5 year bond. From Table 6 we see that with a single bond we no longer get these highly levered positions. Furthermore, the confidence intervals are much narrower, which is due to the more precise estimation of the parameter estimates as well as to the number of parameters used to calculate the strategy. To determine the investment strategy in a three-factor essentially affine model we use 25 different parameters. In the one-factor essentially affine model only 5 different parameters are used.

[Table 6 about here.]

Finally, an investor who bases his investment strategy on a one-factor completely affine model (i.e. $m = 1$ and $\lambda_X = 0$ in (1)-(3)) will follow the investment strategy

$$\pi(X,t) = -\frac{1}{\gamma}\left(\hat{\sigma}_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{\lambda}_0 - \frac{1-\gamma}{\gamma}\left(\hat{\sigma}_X \hat{\mathcal{B}}(\tilde{\tau})\right)^{-1}\hat{H}(\tau). \tag{19}$$

The investor uses the parameter estimates in Panel B in Table 3. The investment strategy is displayed in Table 7 for different investment horizons and the associated 95% confidence intervals calculated by the use of each of the draws from the posterior distribution of the parameter estimates for the one-factor completely affine model. Panel A shows the portfolio weights for an investor with a risk aversion of $\gamma = 5$, while Panel B shows the weights for an investor with a risk aversion of $\gamma = 10$. As in the one-factor essentially

[11] To make the investment strategy depend on the short rate instead of an artificial state variable we can rewrite the one-factor model as follows:

$$dr_t = \kappa\left(\hat{\theta} - r_t\right)dt + \sigma_X \delta_X\, dz_t,$$

where $\hat{\theta} = \delta_X \theta + \delta_0$ and the market price of risk is given by $\lambda = \left(\lambda_0 - \frac{\lambda_X}{\delta_X}\delta_0\right) + \frac{\lambda_X}{\delta_X} r$.

affine model the investor only invests in a single bond and the bank account. For now we assume that the bond is a 5-year-to-maturity bond. From Table 7 we see, as for the one-factor essentially affine model, that with a single bond we no longer get these highly levered positions. Comparing the strategies for the two one-factor models in Tables 6 and 7, we see that initially they are more or less the same. However, the investor following the one-factor essentially affine model takes slightly higher positions in the bond for longer investment horizons. Even though the two strategies look more or less the same initially, this is not necessarily the case as time passes. The reason is that the strategy based on the one-factor essentially affine model is state dependent, which is not the case for the strategy based on the completely affine model.

[Table 7 about here.]

4.2 Utility loss

We will now compare the performance of the four different investment strategies. All four investment strategies will carry a utility loss. The loss is due to parameter uncertainty and for three of the strategies also due to model misspecification. We have just seen that the parameter uncertainty increases with the number of term structure factors and with the complexity of the risk premium. Hence, even if the investor has identified the true data generating model he will suffer a loss due to parameter uncertainty. We will now investigate if an investment strategy based on one of the more parsimonious – but misspecified – models with less parameter uncertainty will carry a lower utility loss.

MCMC is ideal to examine this tradeoff. To see this we note that in the estimation we obtain N draws $\theta_1, ..., \theta_N$ from the distribution of model parameters given the data, $p(\Theta|Y)$. Now, let $M_j[\Theta]$ denote model j with associated parameters Θ. To calculate the expected loss of following a portfolio strategy using model i with associated model parameters $\tilde{\Theta}$ when the true model is model j with associated parameters Θ. We write the loss as $L(M_i[\tilde{\Theta}]|Y, M_j[\Theta])$.[12] To calculate the *average* loss

$$E\big(L(M_i[\tilde{\Theta}]|Y, M_j[\Theta])\big)$$

we note that $E\big(L(M_i[\tilde{\Theta}]|Y, M_j[\Theta])\big) = \int L(M_i[\tilde{\Theta}]|Y, M_j[x])p(x|Y)dx$. The last expression

[12]Specifically, $L(M_j[\Theta]|Y, M_j[\Theta]) = 0$

can be evaluated by averaging the losses across the MCMC draws

$$\int L(M_i[\tilde{\Theta}]|Y, M_j[x])p(x|Y)dx = \sum_{i=1}^{N} L(M_i[\tilde{\Theta}]|Y, M_j[\theta_i]).$$

First we consider the investment strategy based on the true model. Hence, as explained above, to calculate the expected utility loss we use each draw from the posterior distribution as the true parameters and calculate the loss from following the portfolio rule based on the posterior median parameters given in Table 1. The expected utility loss is then found by averaging the losses over all draws. The wide confidence intervals for the investment strategies in Table 4 indicate that an investor using the true model might suffer a considerable utility loss due to the high parameter uncertainty. Panel A in Figure 2 shows the expected utility loss as a function of the investor's investment horizon for four different levels of risk aversion. Intuitively, the wealth loss is increasing in the investor's investment horizon. However, the loss is high even for short horizons. For example, an investor with an investment horizon of 5 years and a risk aversion of 5 will suffer a wealth loss of 61% due to parameter uncertainty. For an investment horizon less than 1 year the loss is decreasing in the investor's risk aversion, i.e the investor with the highest risk aversion suffers the lowest loss. In contrast, for longer horizons, $T > 2$, the investor with the highest risk aversion will suffer the biggest loss, whereas the investor with $\gamma = 2$ will suffer the lowest loss. The longer the investment horizon and the higher the risk aversion, the higher the hedge demand. From Table 4 we see that the allocations in the hedge portfolio are highly levered, and hence using a wrong set of parameter estimates may have a big influence on the hedge strategy and lead to huge utility losses. On the other hand, an investor with a low risk aversion has a high speculative demand and therefore experiences a relatively high loss both for a short and a long investment horizon.

Panel B in Figure 2 display the distribution of the 1,000 MCMC draws for an investor with an investment horizon of $T = 5$ years. Hence, the expected utility loss in Panel A for $T = 5$ equals the mean of the 1,000 losses displayed in Panel B. The distribution is illustrated for four different levels of the investor's risk aversion. For all four levels of risk aversion the investor suffers a loss between 90 and 100% for some of draws. The higher the risk aversion, the higher the probability of suffering a loss in the 90-100% interval. In particular, for an investor with a risk aversion of $\gamma = 10$ almost 60% of the draws imply a loss between 90 and 100%. Hence, for some of the draws the investment strategy based on the true model may induce huge losses because of parameter uncertainty.

[Figure 2 about here.]

Figure 2 clearly shows that just because an investor bases his investments on the true model, he can go very wrong if he does not use the true set of parameter estimates. Remember that these losses are based on parameter estimates estimated by the use of daily data on 6 different year-to-maturity bonds over a period of 35 years, and all sets of parameter estimates satisfy the model-implied parameter restrictions.

Why do we see such big losses? From (11) we see that the market price of risk vector, $\lambda = \lambda_0 + \lambda_X X$, determines the speculative part of the investment. Furthermore, the H- and G-functions, which determine the hedge allocation, are directly affected by λ_0, λ_X, and also indirectly affected by λ_X through $\kappa = \tilde{\kappa} - \lambda_X$.[13] Consequently, it is important to get a good estimate of λ_0 as well as λ_X. However, according to Table 1, λ_0 and λ_X are the two parameters with the widest 95% confidence intervals. Another parameter we would expect to have a big influence on the loss is the κ-matrix. $\kappa = \tilde{\kappa} - \lambda_X$ under $\mathbb{P}$, where $\tilde{\kappa}$ is the speed-of-mean-reversion matrix under the risk-neutral probability measure $\mathbb{Q}$. It is a well-known fact that the parameter estimates under $\mathbb{Q}$ are associated with small confidence intervals, hence by assuming $\hat{\lambda}_X = \lambda_X$ we get that $\hat{\kappa} \approx \kappa$.[14] Panel A in Figure 3 displays the loss an investor with a risk aversion of $\gamma = 5$ suffers if he knows the true value of λ_X but not of any other parameters. For longer horizons we see that the investor experiences a reduction in his wealth loss if he knows the true value of λ_X. For shorter horizons the loss actually increases if the investor knows the true value of λ_X. Also there is no reduction in the expected loss if the investor knows the true value of λ_0 according to Panel B in Figure 3. However, remembering the definition of the market price of risk, $\lambda = \lambda_0 + \lambda_X X$, λ_X determines the trend in the market price while λ_0 determines the level. It seems reasonable that the trend in the market price is more important than the level. As a final exercise we put both $\hat{\lambda}_X = \lambda_X$ and $\hat{\lambda}_0 = \lambda_0$. This implies a huge reduction in the expected loss as illustrated in Panel C in Figure 3. Here we illustrate the expected loss for four different levels of the investor's risk aversion, $\gamma \in \{1, 2, 5, 10\}$. Note that the vertical axis only runs up to 1%. The log-investor, $\gamma = 1$, suffers the biggest loss, but even

[13] $\tilde{\kappa}$ is the speed of mean-reversion matrix under the risk-neutral probability measure defined in (4).

[14] See for example Dai and Singleton (2002) and Duffee (2002).

with an investment horizon of 10 years the loss is only 0.30%. The fact that we see such small losses makes sense given that, when $\hat{\lambda}_X = \lambda_X$ and $\hat{\lambda}_0 = \lambda_0$, the only parameters that differ from our benchmark are δ_0, δ_X and $\tilde{\kappa}$. These three parameters are all associated with quite small confidence intervals compared to λ_0 and λ_X.

[Figure 3 about here.]

The above discussion documents that an investor can go very wrong when not taking parameter uncertainty into account. We will now investigate if an investment strategy based on one of the more parsimonious – but misspecified – models actually induces a lower utility loss because of the lower uncertainty in the parameters of the simpler models. Table 8 displays the utility loss for an investor with $\gamma = 5$ and $T = 5$ under the assumption of known and unknown estimates for the true model, respectively. That is, all the losses are calculated under the assumption that the three-factor essentially affine model is the true model. The losses in the column named "No parameter uncertainty" are calculated under the assumption that the true set of parameter estimates are known and given by Table 1. Obviously, an investor basing his investment on the true model under the assumption of known parameters will suffer a loss of 0%. Consistent with the findings in e.g. Sangvinatsos and Wachter (2005) we find huge utility losses for ignoring the affine risk premia in the term structure under the assumption of known parameter estimates. An investor using the three-factor completely affine with constant risk premia will suffer a loss of almost 49.51%. If instead the investor bases his investment on the one-factor essentially (completely) affine model he will suffer a slightly higher (lower) loss. Hence, under the assumption of known parameters the suboptimal investment strategies induce large costs, so that it appears to be crucial to base the investment strategy on the true model.[15] However, this conclusion is invalid once parameter uncertainty is taken into account. In this case, an investor using the true model will suffer an average loss of 50.59%. On the other hand, an investor basing his investment on the three-factor completely affine model will suffer a slightly lower loss of 59.49%, whereas an investor basing his investment on the one-factor essentially affine model will suffer a loss of 54.55%. However, the investor who will suffer the lowest loss of all is the one basing his investment strategy on the most parsimonious model, the one-

[15]This is consistent with findings in e.g. Sangvinatsos and Wachter (2005), Koijen, Nijman, and Werker (2010), and Larsen and Munk (2012).

factor completely affine model. He will suffer a loss of 50.97%, which is a 20% lower cost compared to the investment strategy based on the true model.

[Table 8 about here.]

Figure 4 displays the expected utility loss as a function of the investor's investment horizon for the four investment strategies. The four panels illustrate the losses for four different levels of risk aversion, $\gamma \in \{1, 2, 5, 10\}$. The expected utility loss from basing the investment on the true model is illustrated with a red line, the expected utility loss from basing the investment on the three-factor completely affine model is illustrated with a green line, whereas the expected utility loss from basing the investment on the one-factor completely (essentially) affine model is illustrated with a cyan (blue) line. Intuitively, the expected utility loss is increasing in the investor's investment horizon for all four models. From Panel A and Panel B it follows that for low levels of risk aversion the investor who bases his investment on the correct model clearly suffers the lowest loss. No matter which of the three more parsimonious models the investor bases his investment strategy on will induce the investor with more less the same level of loss. For $\gamma = 5$ and higher this picture changes. For investment horizons exceeding two years, it is now the strategy based on the most parsimonious model that produces the lowest expected loss to the investor. With a risk aversion of $\gamma = 5$, the strategies based on the two three-factor models imply more or less the same cost for the investor and are the two strategies inducing the highest cost. The one-factor essentially affine model induces a slightly smaller loss, whereas the one-factor completely affine model induces the lowest loss of all four models. For $\gamma = 10$, the one-factor completely affine model clearly outperforms the other three models. The strategy based on the correct model gives the worst performance of all four models except for investment horizons less than one year. In particular, under the assumption of a risk aversion of $\gamma = 10$ and an investment horizon of 5 years, an investor basing his investment strategy on the true model will suffer a loss of 66%, whereas an investor basing his investment strategy on the most parsimonious – and misspecified – model only suffers a loss of 34%, a loss reduction of approximately 48%.

[Figure 4 about here.]

The above discussions are based on the expected utility loss based on the 1,000 MCMC draws. Remember that to calculate the expected utility loss, we use each draw from the posterior distribution as the true parameters and calculate the loss from following the portfolio rule based on the posterior median parameters given for each of the four models. The expected utility is then found by averaging the losses over all draws. Figure 5 displays the distribution of the 1,000 losses determined by the use of the 1,000 draws from the posterior distribution. The four panels show the distribution for the four different investment strategies. In each panel the investment horizon is 5 years and the loss is displayed for four different levels of the investor's risk aversion. From Panel A we see that an investor using the correct model has a high probability of suffering a loss between 90 and 100% if his risk aversion is five or higher. This is not the case for any of the other three model. On the other hand, for a log investor the probability of a loss between between 90 and 100% is quite low for the true model, but very high for the three misspecified models.

[Figure 5 about here.]

So far, the expected utility loss in the one-factor models has been determined under the assumption that the investor invests in a 5-year bond. As a robustness check, Figure 6 depicts the expected utility loss for an investor having access to a 1-year, 5-year, and 10-year bond, respectively. The four panels illustrates the losses for four different levels of risk aversion. The investor bases his investment on the one-factor completely affine model. Obviously, the expected utility loss is little sensitive to the time-to-maturity of the bond. This holds true for all levels of risk aversion.

[Figure 6 about here.]

5 Conclusion

TO COME!

A Proofs

A.1 Proof of Proposition 1

To solve for the optimal investment strategy we use the Dynamic Programming Approach suggested by Merton (1969, 1971, 1973). The Hamilton-Jacobi-Bellman (HJB) equation associated with the dynamic optimization problem is given by

$$
\begin{aligned}
0 = \sup_{\pi \in \mathbb{R}^d} \Big\{ & J_t + J_W \left[r - \pi' \mathcal{B}(\tilde{\tau})' \sigma_X \lambda \right] W + \frac{1}{2} J_{WW} W^2 \pi' \mathcal{B}(\tilde{\tau})' \sigma_X \sigma_X' \mathcal{B}(\tilde{\tau}) \pi \\
& + J_X' \kappa \left(\theta - X \right) + \frac{1}{2} \mathrm{tr} \left(J_{XX}\, \sigma_X\, \sigma_X' \right) - W \pi' \mathcal{B}(\tilde{\tau})' \sigma_X \sigma_X' J_{WX} \Big\},
\end{aligned} \tag{20}
$$

with the terminal condition $J(W, X, T) = \frac{W^{1-\gamma}}{1-\gamma}$ if $\gamma \neq 1$ and $J(W, X, T) = \ln W$ if $\gamma = 1$. The subscripts on J denote the partial derivatives.

The first order condition w.r.t. π implies that a candidate for the optimal investment strategy is given by

$$
\pi^*(W, X, t) = \frac{J_W}{J_{WW} W} \left(\sigma_X' \mathcal{B}(\tilde{\tau}) \right)^{-1} \lambda + \frac{1}{J_{WW} W} \left(\sigma_X' \mathcal{B}(\tilde{\tau}) \right)^{-1} \sigma_X' J_{WX}. \tag{21}
$$

By substituting the candidate for the optimal investment strategy into the HJB-equation we get that

$$
\begin{aligned}
0 = & J_t + J_W \left[r - (\pi^*)' \mathcal{B}(\tilde{\tau})' \sigma_X \eta \right] W + \frac{1}{2} J_{WW} W^2 (\pi^*)' \mathcal{B}(\tilde{\tau})' \sigma_X \sigma_X' \mathcal{B}(\tilde{\tau}) \pi^* \\
& + J_X' \kappa \left(\theta - X \right) + \frac{1}{2} \mathrm{tr} \left(J_{XX}\, \sigma_X\, \sigma_X' \right) - W (\pi^*)' \mathcal{B}(\tilde{\tau})' \sigma_X \sigma_X' J_{WX}.
\end{aligned} \tag{22}
$$

An educated guess of the solution is

$$
J(W, X, t) = \frac{1}{1-\gamma} \left(W e^{F(\tau) + H(\tau)' X + \frac{1}{2} X' G(\tau) X} \right)^{1-\gamma}. \tag{23}
$$

The terminal condition of the HJB-equation implies that $F(0) = H(0) = G(0) = 0$. Substituting the candidate for the optimal investment strategy (21) and the relevant derivatives of our guess into the HJB-equation, simplifying, and finally matching coefficients on $X'[\,\cdot\,]X$, X', and the constant terms lead to the following system of ordinary differential equations:

$$\begin{aligned}\frac{\partial F(\tau)}{\partial t} &= H(\tau)'\left[\kappa\theta + \frac{1-\gamma}{\gamma}\sigma_X\lambda_0\right] + \frac{1-\gamma}{2\gamma}H(\tau)'\sigma_X\sigma_X' H(\tau)\\ &\quad + \frac{1}{4}\mathrm{tr}\left(\left(G(\tau)+G(\tau)'\right)\sigma_X\sigma_X'\right) + \frac{1}{2\gamma}\lambda_0'\lambda_0 + \delta_0\end{aligned} \tag{24}$$

$$\begin{aligned}\frac{\partial H(\tau)}{\partial t} &= \left[\frac{1-\gamma}{\gamma}\lambda_X'\sigma_X' - \kappa' + \frac{1-\gamma}{2\gamma}\left(G(\tau)+G(\tau)'\right)\sigma_X\sigma_X'\right]H(\tau)\\ &\quad + \frac{1}{2}\left(G(\tau)+G(\tau)'\right)\left[\kappa\theta + \frac{1-\gamma}{\gamma}\sigma_X\lambda_0\right] + \frac{1}{\gamma}\lambda_X'\lambda_0 + \delta_X\end{aligned} \tag{25}$$

$$\begin{aligned}\frac{\partial G(\tau)}{\partial t} &= \frac{1}{\gamma}\lambda_X'\lambda_X + \frac{1-\gamma}{\gamma}\lambda_X'\sigma_X'\left(G(\tau)+G(\tau)'\right) - \left(G(\tau)+G(\tau)'\right)\kappa\\ &\quad + \frac{1-\gamma}{4\gamma}\left(G(\tau)+G(\tau)'\right)\sigma_X\sigma_X'\left(G(\tau)+G(\tau)'\right).\end{aligned} \tag{26}$$

Hence, our guess (23) is the solution to the HJB-equation if F, H, and G solve the above system of ODE's. Finally, substituting the relevant derivatives of J into (21) gives the optimal strategy (11).

A.2 Proof of Proposition 2

Any combination of an initial wealth and an investment strategy π will give rise to a terminal wealth W_T^π and the expected utility associated with that is thus given by

$$\hat{J}(W,X,t) = \begin{cases} \mathbb{E}_t^{\mathbb{P}}\left[\frac{1}{1-\gamma}\left(W_T^\pi\right)^{1-\gamma}\right], & \gamma > 1,\\ \mathbb{E}_t^{\mathbb{P}}\left[\ln\left(W_T^\pi\right)\right], & \gamma = 1.\end{cases} \tag{27}$$

From Theorem 2 in Larsen and Munk (2012) it follows that the expected utility generated by the investment strategy, π, is given by

$$\hat{J}(W,X,t) = \begin{cases} \frac{1}{1-\gamma}\left(W\, e^{C(X,T-t)}\right)^{1-\gamma}, & \gamma > 1,\\ \ln W + C(X,T-t), & \gamma = 1,\end{cases} \tag{28}$$

where the function $C(X, T-t)$ satisfies the PDE

$$\begin{aligned}&\frac{\partial C}{\partial t} + \left(\kappa\left(\theta - X\right) - (\gamma-1)\,\sigma_X\sigma_P(\tilde{\tau})'\pi(X,t)\right)'\frac{\partial C}{\partial X} + \frac{1}{2}\,\mathrm{tr}\left(\frac{\partial^2 C}{\partial X^2}\sigma_X\sigma_X'\right)\\ &- \frac{\gamma-1}{2}\left(\frac{\partial C}{\partial X}\right)'\sigma_X\sigma_X'\frac{\partial C}{\partial X} + r(X) + \pi(X,t)'\sigma_P(\tilde{\tau})\left[\lambda(X) - \frac{\gamma}{2}\sigma_P(\tilde{\tau})'\pi(X,t)\right] = 0\end{aligned} \tag{29}$$

with the terminal condition $C(X,0) = 0$. $\sigma_P(\tilde{\tau})$ denotes the $n \times m$ volatility-matrix of the traded zero-coupon bonds. For the specific suboptimal investment strategy given in (12) an educated guess on a solution to the PDE is given by[16]

$$C(X,t) = C_1(\tau) + C_2(\tau)'X + \frac{1}{2}X'C_3(\tau)X.$$

Substituting in the relevant derivatives, the relevant investment strategy in (12), simplifying, and finally matching coefficients on $X'[\cdot]X$, X', and the constant terms leads to the following system of ordinary differential equations:

$$\begin{aligned}
\frac{\partial C_1}{\partial t} &= C_2(\tau)' \left[\kappa\theta + \frac{\gamma-1}{\gamma}\sigma_X\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0\right] - \frac{\gamma-1}{2}C_2(\tau)'\sigma_X\sigma_X'C_2(\tau) \\
&+ \frac{\gamma-1}{\gamma}\hat{\hat{H}}(\tau)'\hat{\sigma}_X\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\left[\lambda_0 + \sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0 - (\gamma-1)\,\sigma_X'C_2(\tau)\right] \\
&- \frac{(\gamma-1)^2}{2\gamma}\hat{\hat{H}}(\tau)'\hat{\sigma}_X\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\sigma}_X'\hat{\hat{H}}(\tau) \\
&+ \frac{1}{4}\mathrm{tr}\left(\left(C_3(\tau)+C_3(\tau)'\right)\sigma_X\sigma_X'\right) - \frac{1}{\gamma}\hat{\hat{\lambda}}_0'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\lambda_0 \\
&- \frac{1}{2\gamma}\hat{\hat{\lambda}}_0'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0 + \delta_0
\end{aligned} \tag{30}$$

$$\begin{aligned}
\frac{\partial C_2}{\partial t} &= \left[\frac{\gamma-1}{\gamma}\hat{\hat{\lambda}}_X'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\sigma_X' - \kappa' - \frac{\gamma-1}{2}\left(C_3(\tau)+C_3(\tau)'\right)\sigma_X\sigma_X'\right]C_2(\tau) \\
&+ \frac{1}{2}\left(C_3(\tau)+C_3(\tau)'\right)\left[\kappa\theta + \frac{\gamma-1}{\gamma}\sigma_X\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\left(\hat{\hat{\lambda}}_0 - (\gamma-1)\,\hat{\sigma}_X'\hat{\hat{H}}(\tau)\right)\right] \\
&+ \frac{\gamma-1}{\gamma}\left[\lambda_X'\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\sigma}_X' + \hat{\hat{\lambda}}_X'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\sigma}_X'\right]\hat{\hat{H}}(\tau) \\
&+ \frac{\gamma-1}{2\gamma}\left(\hat{\hat{G}}(\tau)+\hat{\hat{G}}(\tau)'\right)\hat{\sigma}_X\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\left[\lambda_0 + \sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0\right] \\
&- \frac{(\gamma-1)^2}{2\gamma}\left(\hat{\hat{G}}(\tau)+\hat{\hat{G}}(\tau)'\right)\hat{\sigma}_X\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\left[\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\sigma}_X'\hat{\hat{H}}(\tau) + \sigma_X'C_2(\tau)\right] \\
&- \frac{1}{\gamma}\lambda_X'\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0 - \frac{1}{\gamma}\hat{\hat{\lambda}}_X'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\lambda_0 \\
&- \frac{1}{\gamma}\hat{\hat{\lambda}}_X'\left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1}\sigma_P(\tau)\sigma_P(\tau)'\left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1}\hat{\hat{\lambda}}_0 + \delta_X
\end{aligned} \tag{31}$$

[16]For the completely affine models the matrix function $C_3(\cdot)$ is put equal to zero.

$$
\begin{aligned}
\frac{\partial C_3(\tau)}{\partial t} &= -\frac{1}{\gamma}\hat{\hat{\lambda}}_X' \left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1} \sigma_P(\tau) \left(2\lambda_X + \sigma_P(\tau)' \left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1} \hat{\hat{\lambda}}_X\right) \\
&+ \frac{\gamma-1}{\gamma}\lambda_X'\sigma_P(\tau)' \left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1} \hat{\sigma}_X \left(\hat{\hat{G}}(\tau) + \hat{\hat{G}}(\tau)'\right) \\
&+ \frac{\gamma-1}{\gamma}\hat{\hat{\lambda}}_X' \left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1} \sigma_P(\tau)\sigma_X' \left(C_3(\tau) + C_3(\tau)'\right) \\
&+ \frac{\gamma-1}{\gamma}\hat{\hat{\lambda}}_X' \left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1} \sigma_P(\tau)\sigma_P(\tau)' \left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1} \hat{\sigma}_X' \left(\hat{\hat{G}}(\tau) + \hat{\hat{G}}(\tau)'\right) \\
&- \left(C_3(\tau) + C_3(\tau)'\right)\kappa - \frac{\gamma-1}{4}\left(C_3(\tau) + C_3(\tau)'\right)\sigma_X\sigma_X' \left(C_3(\tau) + C_3(\tau)'\right) \\
&- \frac{(\gamma-1)^2}{4\gamma}\left(\hat{\hat{G}}(\tau) + \hat{\hat{G}}(\tau)'\right)\hat{\sigma}_X \left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1} \sigma_P(\tau)\sigma_P(\tau)' \left(\hat{\sigma}_X'\hat{B}(\tilde{\tau})\right)^{-1} \hat{\sigma}_X' \left(G(\tau) + G(\tau)'\right) \\
&- \frac{(\gamma-1)^2}{2\gamma}\left(\hat{\hat{G}}(\tau) + \hat{\hat{G}}(\tau)'\right)\hat{\sigma}_X \left(\hat{B}(\tilde{\tau})'\hat{\sigma}_X\right)^{-1} \sigma_P(\tau)\sigma_X' \left(C_3(\tau) + C_3(\tau)'\right).
\end{aligned} \tag{32}
$$

with boundary condition $C_1(0) = C_2(0) = C_3(0) = 0$. The hats ($\hat{}$) on the parameters imply that the benchmark parameter estimates for the models should be used. For the three-factor essentially affine model as well as for the two completely affine models $\hat{\hat{\lambda}}_X = \hat{\lambda}_X$, $\hat{\hat{\lambda}}_0 = \hat{\lambda}_0$, $\hat{\hat{H}}(\cdot) = \hat{H}(\cdot)$, and $\hat{\hat{G}}(\cdot) = \hat{G}(\cdot)$, that is use the benchmark estimates for the three investment strategies. For the one-factor essentially affine model

$$
\begin{aligned}
\hat{\hat{\lambda}}_X &= \frac{\hat{\lambda}_X}{\hat{\delta}_X}\delta_X', \qquad & \hat{\hat{\lambda}}_0 &= \hat{\lambda}_0 - \frac{\hat{\lambda}_X}{\hat{\delta}_X}\left(\hat{\delta}_0 - \delta_0\right), \\
\hat{\hat{G}}(\tau) &= \frac{\hat{G}(\tau)}{\hat{\delta}_X}\delta_X', \qquad & \hat{\hat{H}}(\tau) &= \hat{H}(\tau) - \frac{\hat{G}(\tau)}{\hat{\delta}_X}\left(\hat{\delta}_0 - \delta_0\right).
\end{aligned}
$$

The functions $\hat{H}(\tau)$ and $\hat{G}(\tau)$ solve the system of ordinary differential equations (25)-(26) from the optimal setup with the benchmark parameter estimates for each of the four models. Hence, our guess is the solution to the PDE (29) if C_1, C_2, and C_3 solve the above system of ODE's.[17]

B Details on the MCMC estimation

First the conditionals mentioned in the text are derived, and thereafter practical issues regarding the MCMC sampler are discussed.

[17]Note for the completely affine models we have that $G(\cdot) = C_3(\cdot) = 0$ and hence the system of ordinary differential equations can be simplified a lot.

B.1 Conditional Distributions

B.1.1 The Conditionals $p(X|\Theta)$ and $p(Y|\Theta, X)$

The conditional $p(X|\Theta)$ is used in several steps of the MCMC procedure and is calculated as

$$p(X|\Theta) = \left(\prod_{t=1}^{T} p(X_t|X_{t-1}, \Theta) \right) p(X_0).$$

The continuous-time specification in (4) is approximated using an Euler scheme[18]

$$X_{t+1} = X_t + \mu_{Xt}^{\mathbb{P}} \Delta_t + \sqrt{\Delta_t}\, \xi_{t+1},$$

where $\xi_{t+1} \sim N(0, I_N)$, Δt is the time between two observations, and $\mu_{Xt}^{\mathbb{P}}$ is the drift under $\mathbb{P}$. Therefore

$$p(X|\Theta) \propto \exp\left\{ -\frac{1}{2\Delta_t} \sum_{t=1}^{T} \sum_{i=1}^{3} \left[X_t - X_{t-1} - \mu_{Xt-1}^{\mathbb{P}} \Delta_t \right]_i^2 \right\} p(X_0).$$

If the difference between the actual yields and the model-implied yields at time t is denoted by $\hat{e}_t = Y_t - (A(\Theta) + B(\Theta)X_t)$, the density $p(y|\Theta, X)$ can be written as

$$p(Y|\Theta, X) \propto \prod_{i=1}^{k} \left(D_{ii}^{-\frac{T}{2}} \exp\left\{ -\frac{1}{2D_{ii}} \sum_{t=1}^{T} \hat{e}_{t,i}^2 \right\} \right) \propto \varphi^{-kT} \exp\left\{ -\frac{1}{2\varphi^2} \sum_{t=1}^{T} \hat{e}_t' \hat{e}_t \right\}.$$

B.1.2 The Hybrid MCMC algorithm

According to Bayes' theorem the conditional of the risk premium parameters is given as

$$\begin{aligned} p(\lambda_0, \lambda_X|\Theta_{\backslash\lambda_0,\lambda_X}, X, Y) &\propto p(Y|\Theta, X)\, p(\lambda_0, \lambda_X|\Theta_{\backslash\lambda_0,\lambda_X}, X) \\ &\propto p(X|\Theta)\, p(\lambda_0, \lambda_X|\Theta_{\backslash\lambda_0,\lambda_X}), \end{aligned}$$

where $\Theta_{\backslash\lambda_0,\lambda_X}$ denotes the parameter vector without the parameters λ_0 and λ_X. We assume that the priors are a priori independent and in order to let the data dominate the results a standard diffuse, noninformative prior is adopted so $p(\lambda_0, \lambda_X|\Theta_{\backslash\lambda_0,\lambda_X}, X, Y) \propto$

[18]The Euler scheme introduces some discretization error which may induce bias in the parameter estimates. This possible bias can be reduced using Tanner and Wong (1987)'s data augmentation scheme. However, the discretization bias is likely to be small for daily data.

$p(X|\Theta)$ and λ_0, λ_X can be Gibbs sampled one column at a time from a multivariate normal distribution. The conditionals of the other model parameters are given as

$$p(\Theta_j|\Theta_{\backslash\Theta_j}, X, Y) \propto p(Y|\Theta, X)\, p(\Theta_j|\Theta_{\backslash\Theta_j}, X) \propto p(Y|\Theta, X)\, p(X|\Theta). \tag{33}$$

Equation (33) implies that the conditional of the variance of the measurement errors is given as

$$p(D|\Theta_{\backslash D}, X, Y) \propto p(Y|\Theta, X)\, p(D|\Theta_{\backslash D}).$$

The parameter φ^2 can therefore be Gibbs sampled from the inverse Wishart distribution, $\varphi^2 \sim IW(\sum_{t=1}^{T} \hat{e}_t' \hat{e}_t, kT)$.

To sample $\tilde{\kappa}, \delta_0$, and δ_X we use the Random Walk Metropolis-Hastings algorithm (RW-MH). Equation (33) gives the general expression for the conditional distribution.

The latent processes are sampled by sampling X_t, $t = 0, ..., T$ one at a time using the RW-MH procedure. For $t = 1, ..., T-1$ the conditional of X_t is given as

$$\begin{aligned} p(X_t|X_{\backslash t}, \Theta, Y) &\propto p(X_t|X_{t-1}, X_{t+1}, \Theta, Y_t) \\ &\propto p(Y_t|X_t, \Theta)\, p(X_t|X_{t-1}, X_{t+1}, \Theta) \\ &\propto p(Y_t|X_t, \Theta)\, p(X_t|X_{t-1}, \Theta)\, p(X_{t+1}|X_t, \Theta). \end{aligned}$$

For $t = 0$ the conditional is

$$p(X_0|X_1, \Theta, Y) \propto p(X_1|X_0, \Theta, Y)\, p(X_0) \propto p(X_1|X_0, \Theta)\, p(X_0),$$

while for $t = T$ the conditional is

$$\begin{aligned} p(X_T|X_{\backslash X_T}, \Theta, Y) &\propto p(X_T|X_{T-1}, \Theta, Y) \\ &\propto p(Y_T|X_T, X_{T-1}, \Theta, Y_{\backslash Y_T})\, p(X_T|X_{T-1}, \Theta, Y_{\backslash Y_T}) \\ &\propto p(Y_T|X_T, \Theta)\, p(X_T|X_{T-1}, \Theta). \end{aligned}$$

The efficiency of the RW-MH algorithm depends crucially on the variance of the proposal normal distribution. If the variance is too low, the Markov chain will accept nearly every draw and converge very slowly, while it will reject a too high portion of the draws if the variance is too high. We therefore do an algorithm calibration and adjust the variance in the first eight million draws in the MCMC algorithm. Within each parameter block the variance of the individual parameters is the same, while across parameter blocks the variance may be different. Roberts et al. (1997) recommend acceptance rates close to $\frac{1}{4}$ for models of high dimension and therefore the standard deviation during the algorithm calibration is chosen as follows: Every 100'th draw the acceptance ratio of the parameters in a block is evaluated. If it is less than 5 % the standard deviation is doubled, while if it is more than 40 % it is cut in half. This step is prior to the burn-in period since the convergence results of RW-MH only apply if the variance is constant (otherwise the Markov property of the chain is lost). In estimating each model we use an algorithm calibration period of 3 million draws, where the variances of the normal proposal distributions are set, a burn-in period of 5 million draws, and an estimation period of 5 million draws. We keep every 5,000'th draw in the estimation period, which leaves 1,000 draws. For each parameter, we report point estimates along with univariate confidence bands based on the 2.5% and the 97.5% percentiles of the MCMC draws of the posterior distribution. As in Collin-Dufresne, Goldstein, and Jones (2008) we find point estimates as follows. Let ϕ_i denote the ith parameter draw and $\tilde{\phi}_i$ denote the same vector normalized by the posterior standard deviations. The point estimate is the draw i minimizing:

$$\sum_j \left| \tilde{\phi}_j - \tilde{\phi}_i \right| .$$

This version of the multivariate posterior median ensures that parameter restrictions are satisfied for our parameter estimates, which might not be the case if the point estimates are based on univariate medians.

All random numbers in the estimation are draws from Matlab 7.0's generator which is based on Marsaglia and Zaman (1991)'s algorithm. The generator has a period of almost 2^{1430} and therefore the number of random draws in the estimation is not anywhere near the period of the random number generator.

References

Ang, A., S. Dong, and M. Piazzesi (2007). No-arbitrage taylor rules. Working paper.

Barberis, N. (2000). Investing for the Long Run when Returns are Predictable. *The Journal of Finance 55*, 225–264.

Barry, C. B. and S. J. Brown (1985). Differential Information and Securtiy Market Equilibrium. *Journal of Financial and Quantitative Analysis 20*, 407–422.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Association Series B 36*, 192–236.

Brennan, M. J. (1998). The Role of Learning in Dynamic Portfolio Decisions. *European Finance Review 1*(3), 295–306.

Brennan, M. J. and Y. Xia (2000). Stochastic Interest Rates and the Bond-Stock Mix. *European Finance Review 4*(2), 197–210.

Brennan, M. J. and Y. Xia (2002). Dynamic Asset Allocation under Inflation. *The Journal of Finance 57*, 1201–1238.

Campbell, J. Y. and L. M. Viceira (2001). Who Should Buy Long-Term Bonds? *American Economic Review 91*, 99–127.

Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones (2008). Identification of Maximal Affine Term Structure Models. *The Journal of Finance 63*, 743–795.

Dai, Q. and K. Singleton (2000). Specification Analysis of Affine Term Structure Models. *The Journal of Finance 55*, 1943–1978.

Dai, Q. and K. Singleton (2002). Expectation Puzzles, Time-varying Risk Premia, and Dynamic Models of the Term Structure. *Journal of Financial Economics 3*, 415–441.

Duffee, G. R. (2002). Term Premia and Interest Rate Forecast in Affine Models. *The Journal of Finance 57*, 405–443.

Duffie, D. and R. Kan (1996). A Yield-factor Model of Interest Rates. *Mathematical Finance 6*, 379–406.

Eraker, B. (2001). MCMC Analysis of Diffusion Models with Application to Finance. *Journal of Business and Economic Statistics 19*, 177–191.

Feldhütter, P. (2008). Can Affine Models Match the Moments in Bond Yields? Working paper.

Gelfand, A. E., A. F. Smith, and T.-M. Lee (1992). Bayesian Analysis of Constrained Parameter and Truncated Data problems Using Gibbs Sampling. *Journal of the American Statistical Association 87*, 523–532.

Geman, S. and D. Geman (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Trans. on Pattern Analysis and Machine Intelligence 6*, 721–741.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2006). The U.S. Treasury Yield Curve: 1961 to the Present. Working paper, Divisions of Reasearch and Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Hammersley, J. and P. Clifford (1970). Markov Fields on Finite Graphs and Lattices. Unpublished.

Johannes, M., A. Korteweg, and N. Polson (2011). Sequentil Learning, Predictive Regressions, and Optimal Portfolio Returns. Working paper.

Johannes, M. and N. Polson (2006). MCMC Methods for Financial Econometrics. *Handbook of Financial Econometrics*.

Kaminska, I., D. Vayanos, and G. Zinna (2011). Preferred-Habitat Investors and the US Term Structure of Real Rates. Working paper.

Kim, T. S. and E. Omberg (1996). Dynamic Nonmyopic Portfolio Behavior. *The Review of Financial Studies 9*, 141–161.

Klein, R. W. and V. S. Bawa (1976). The Effect of Estimation Risk on Optimal Portfolio Choice. *Journal of Financial Economics 3*, 215–231.

Koijen, R. S. J., T. Nijman, and B. J. M. Werker (2010). When Can Life-Cycle Investors Benefit from Time-Varying Bond Risk Premia? *The Review of Financial Studies 23*, 741–780.

Korn, R. and H. Kraft (2004). On the Stability of Continuous-Time Portfolio Problems with Stochastic Opportunity Set. *Mathematical Finance 14*, 403–414.

Korteweg, A. G. and N. Polson (2009). Corporate Credit Spreads under Parameter Uncertainty. Working paper.

Larsen, L. S. and C. Munk (2012). The Costs of Suboptimal Dynamic Asset Allocation: General Results and Applications to Interest Rate Risk, Stock Volatility Risk and Growth/Value Tilts. *Journal of Economic Dynamics and Control 36*, 266–293.

Maenhout, P. J. (2004). Robust Portfolio Rules and Asset Pricing. *The Review of Financial Studies 17*, 951–983.

Maenhout, P. J. (2006). Robust Portfolio Rules and Detection-Error Probabilities for a Mean-Reverting Risk Premium. *Journal of Economic Theory 128*, 136–163.

Marsaglia, G. and A. Zaman (1991). A New Class of Random Number Generators. *Annals of Applied Probability 3*, 462–480.

Merton, R. C. (1969). Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case. *Rewiew of Economics and Statistics 51*, 247–257.

Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-time Model. *Journal of Economic Theory 3*, 373–413.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica 41*, 867–887.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953). Equations of State Calculations by fast Computing Machines. *Journal of Chemical Physics 21*, 1087–1091.

Munk, C. and C. Sørensen (2004). Optimal Consumption and Investment Strategies with Stochastic Interest Rates. *Journal of Banking and Finance 28*(8), 1987–2013.

Robert, C. P. and G. Casella (2004). *Monte Carlo Statistical Methods* (Second ed.). Springer-Verlag.

Roberts, G., A. Gelman, and W. Gilks (1997). Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms. *Annals of Applied Probability 7*, 110–120.

Sangvinatsos, A. and J. A. Wachter (2005). Does the Failure of the Expectations Hypotheses Matter for Long-Term Investors? *The Journal of Finance 60*, 179–230.

Sarno, L., P. Schneider, and C. Wagner (2011). Properties of Foreign Exchange Risk Premiums. Journal of Financial Economics, forthcoming.

Sørensen, C. (1999). Dynamic Asset Allocation and Fixed Income Management. *Journal of Financial and Quantitative Analysis 34*, 513–531.

Tanner, M. and W. Wong (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association 82*, 528–549.

Wachter, J. A. (2003). Risk Aversion and Allocation to Long-Term Bonds. *Journal of Economic Theory 112*, 325–333.

Xia, Y. (2001). Learning about Predictatbility: The Effects of Parameter Uncertainty on Dynamic Asset Allocation. *The Journal of Finance 56*, 205–244.

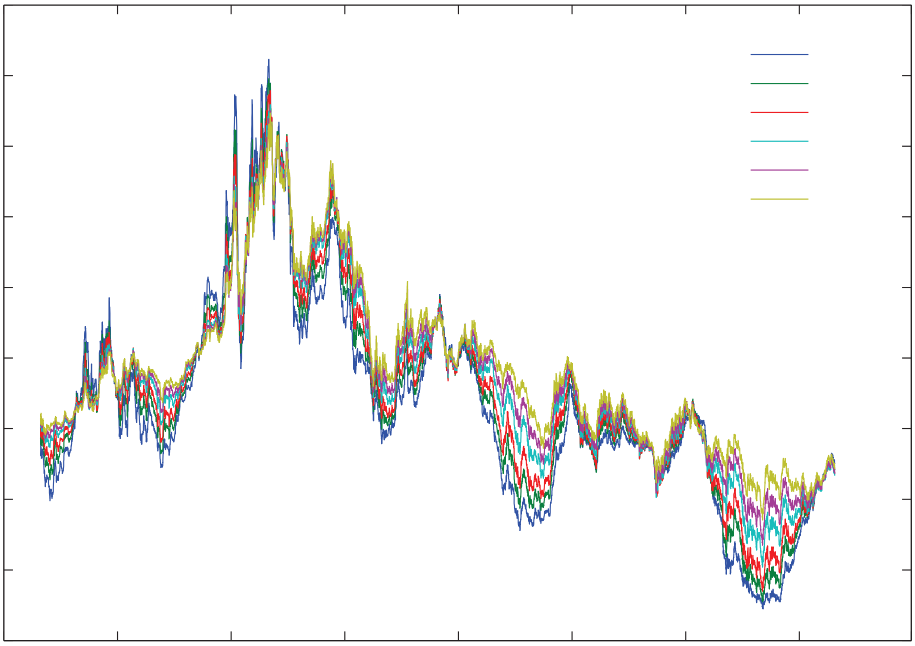

Figure 1: **Time-series of bond yields.** The figure shows daily (continuously compounded) 1, 2, 3, 5, 7, and 10-year zero-coupon yields from August 16, 1971 to August 21, 2006. Source: http://www.federalreserve.gov/pubs/feds/2006. ANDERS GIDER DU LAVE EN NY FIGUR?

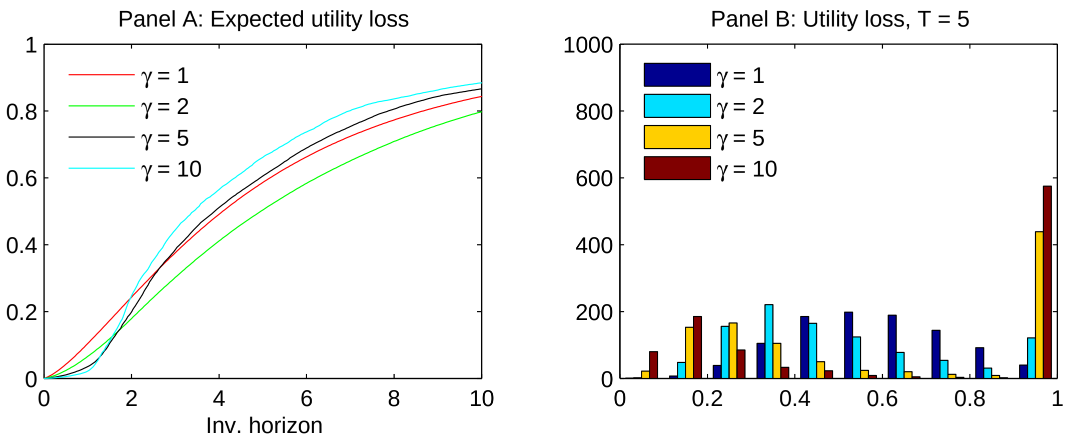

Figure 2: **The utility loss from basing the investment strategy on the three-factor essentially affine model.** Panel A display the expected utility loss as a function of the investor's investment horizon. The loss is displayed for four different levels of the investor's risk aversion, γ. Panel B shows the distribution of the utility loss for an investor with an investment horizon of $T = 5$ years based on the 1,000 MCMC draws. The distribution is illustrated for four different levels of the investor's risk aversion, γ.

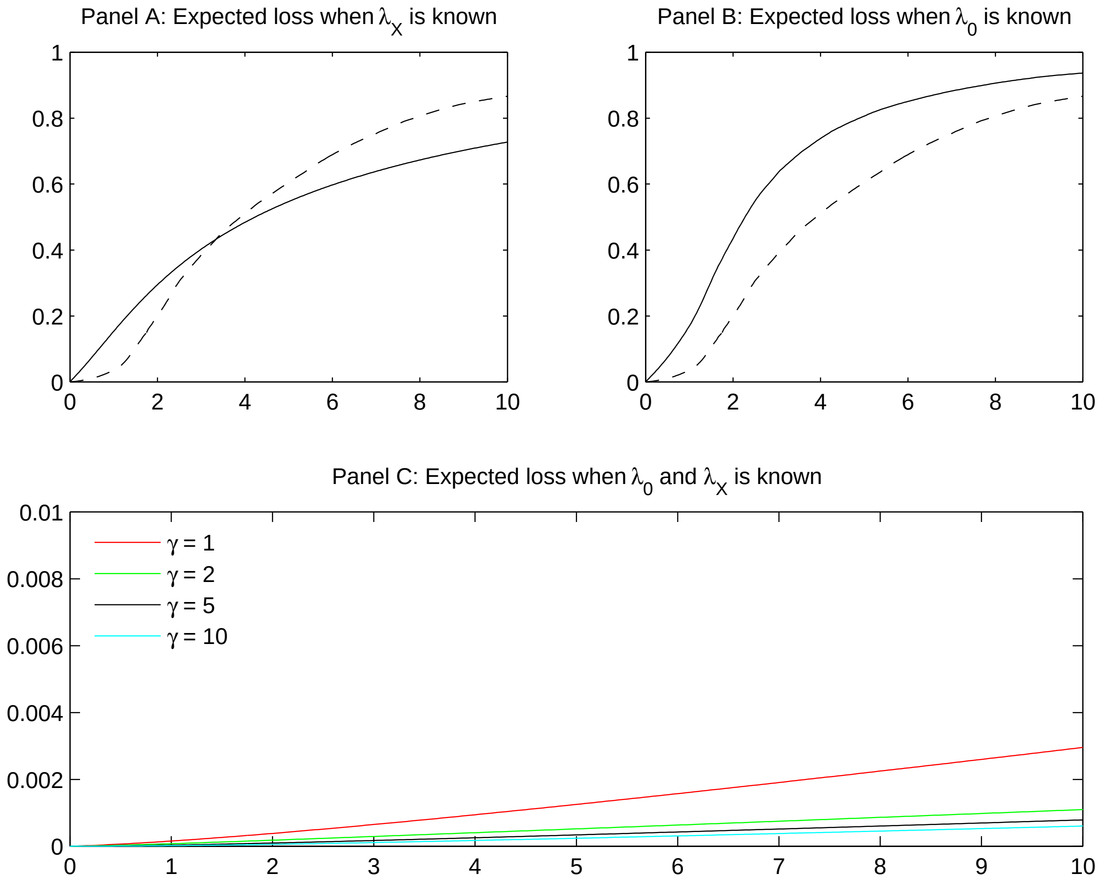

Figure 3: **The utility loss as a function of the investor's investment horizon.** In all three panels the investor bases his investment strategy on the true model, the three-factor essentially affine model. The solid line in Panel A displays the loss for an investor who knows the true value of λ_X, i.e. $\hat{\lambda}_X = \lambda_X$. The solid line in Panel B displays the loss for an investor who knows the true value of λ_0, i.e. $\hat{\lambda}_0 = \lambda_0$. In both panels the investor is assumed to have a risk aversion of $\gamma = 5$, and the dashed line displays the loss when λ_X and λ_0 is unknown. Panel C displays the loss in the case where both λ_X and λ_0 is known for four different values of the investor's risk aversion.

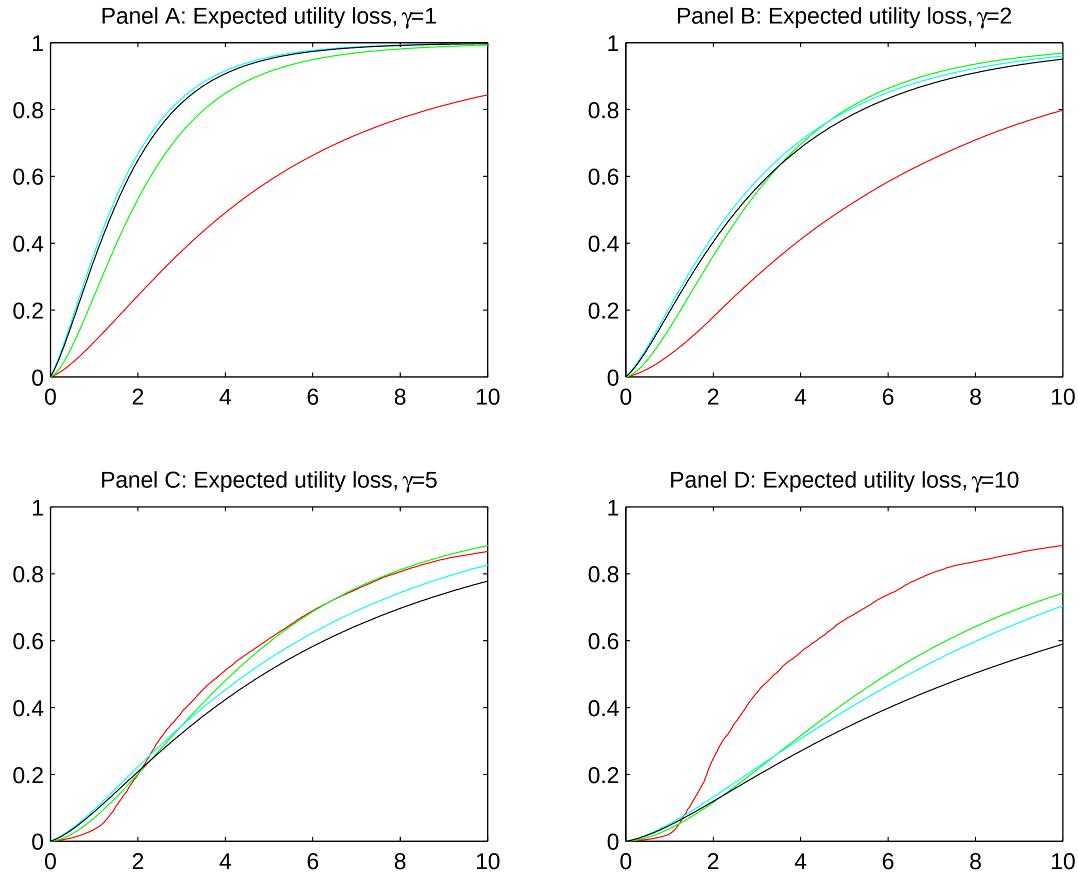

Figure 4: **The expected wealth loss as a function of the investment horizon.** The red line displays the expected utility loss due to parameter uncertainty in the true model, the three-factor essentially affine model. The green line displays the expected loss from basing the investment strategy on the three-factor completely affine model. The cyan line displays the expected loss from basing the investment strategy on the one-factor essentially affine model. The black line displays the expected loss from basing the investment strategy on the one-factor completely affine model. The four panels show the losses for four different levels of the investor's risk aversion.

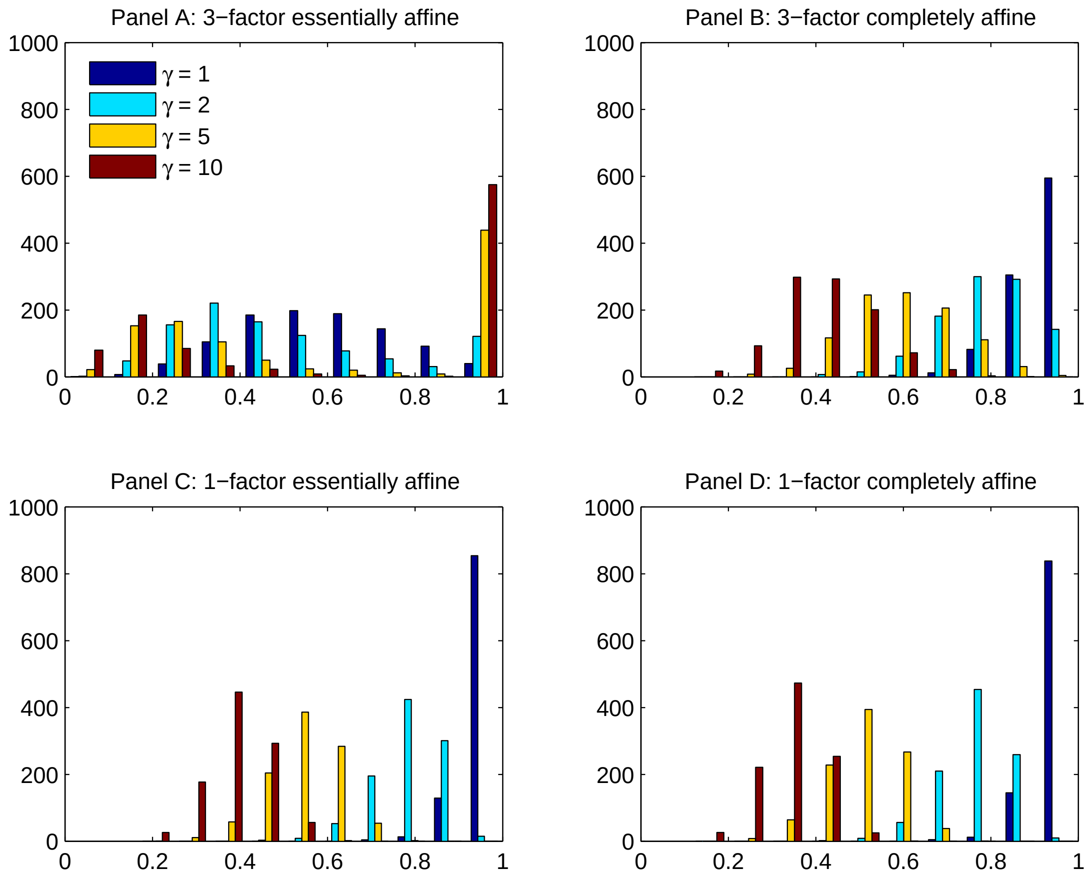

Figure 5: **The distribution of the losses based on the 1,000 MCMC draws.** The distribution of the 1,000 losses determined by the use of the draws from the posterior distribution of the three-factor essentially affine model. The four panels show the distribution for the investment strategies based on the four different term structure models: the true model, the three-factor completely affine model, and the two one-factors models with constant and affine risk premia, respectively. In each panel the loss is displayed for four different levels of the investor's risk aversion. The investor is assumed to have an investment horizon of $T = 5$.

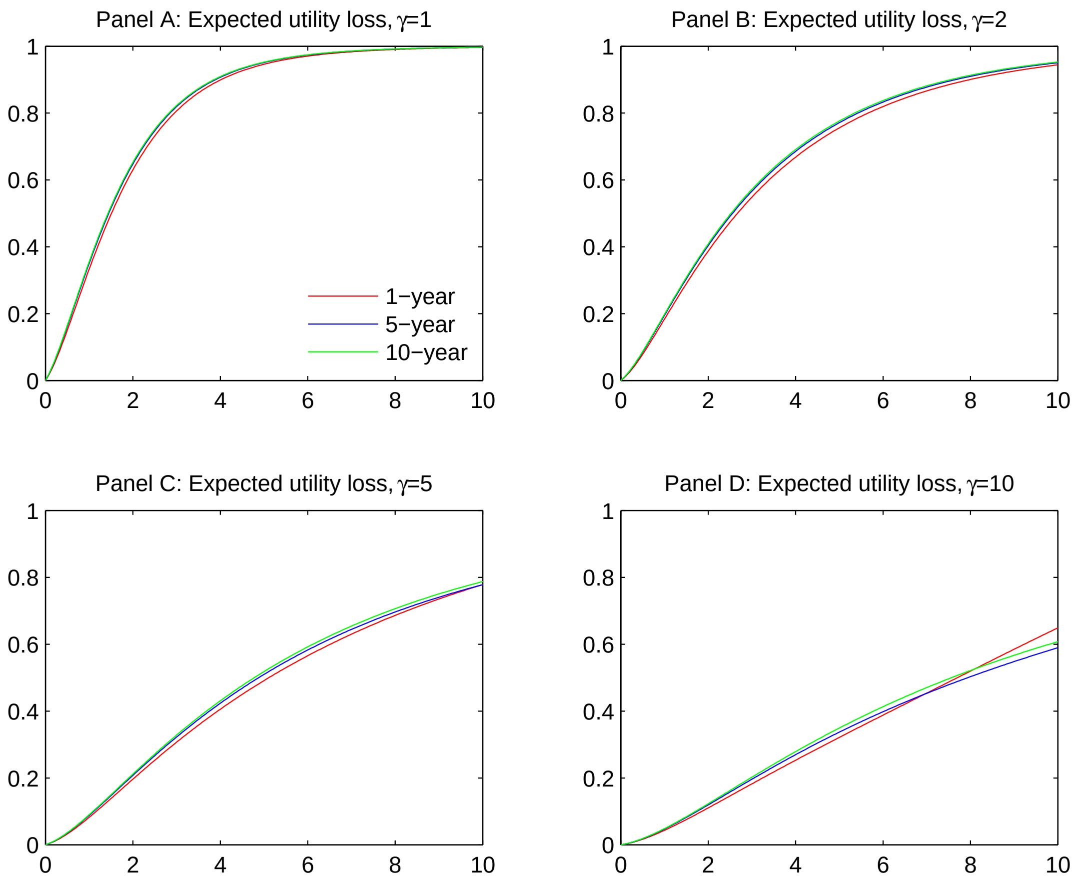

Figure 6: **The expected utility loss as a function of the investor's investment horizon.** All three lines show the loss an investor suffers from basing his investment strategy on the one-factor completely affine model. The red line displays the loss for an investor investing in a 1-year-to-maturity bond. The blue line displays the loss for an investor investing in a 5-year-to-maturity bond. The green line displays the loss for an investor investing in a 10-year-to-maturity bond. The four panels show the losses for four different levels of risk aversion.

Panel A: Scalars			
Parameter	δ_0		
	0.1589		
	$(0.1582; 0.1629)$		

Panel B: Vectors and matrixes			
	Index number i		
Parameter	1	2	3
δ_{Xi}	0.0027	0.0101	0.0101
	$(0.0023; 0.0031)$	$(0.0099; 0.0106)$	$(0.0100; 0.0103)$
κ_{1i}	1.5482	0.1433	0.0271
	$(0.7776; 1.9290)$	$(-0.1706; 0.3661)$	$(-0.0760; 0.1824)$
κ_{2i}	1.5041	0.4016	0.0709
	$(0.5681; 1.7210)$	$(0.2052; 0.7384)$	$(-0.0693; 0.1990)$
κ_{3i}	0.6401	-0.0272	0.1171
	$(0.0553; 1.2182)$	$(-0.3259; 0.2367)$	$(0.0200; 0.2727)$
λ_{0i}	0.1644	-0.4501	-1.081
	$(-1.6999; 0.7431)$	$(-2.0522; 0.4381)$	$(-2.4085; -0.1181)$
λ_{X1i}	-1.1519	-0.1433	-0.0271
	$(-1.5380; 0.3850)$	$(-0.3670; 0.1702)$	$(-0.1829; 0.0760)$
λ_{X2i}	-0.3433	0.4220	-0.0709
	$(-0.5785; 0.5767)$	$(0.0853; 0.6191)$	$(-0.1995; 0.0692)$
λ_{X3i}	-0.1961	0.4645	-0.0979
	$(-0.7848; 0.3685)$	$(0.1990; 0.7546)$	$(-0.2558; -0.0011)$

Table 1: **Parameter estimates for the three-factor essentially affine model.** The model is estimated using daily data on zero-coupon yields from 1971 to 2006. 95%-confidence intervals are given in parentheses. Parameter values are annual and in natural units.

Panel A: Scalars			
Parameter	δ_0		
	0.1826		
	$(0.1809; 0.1838)$		
Panel B: Vectors and matrixes			
	Index number i		
Parameter	1	2	3
δ_{Xi}	0.0027	0.0101	0.0104
	$(0.0023; 0.0031)$	$(0.0098; 0.0104)$	$(0.0101; 0.0105)$
κ_{1i}	0.3984	0	0
	$(0.3925; 0.4055)$	–	–
κ_{2i}	1.1272	0.8179	0
	$(1.0928; 1.1620)$	$(0.8005; 0.8252)$	–
κ_{3i}	0.4348	0.4367	0.0156
	$(0.4259; 0.4695)$	$(0.4250; 0.4562)$	$(0.0153; 0.0158)$
λ_i	-0.0711	-0.4697	-0.4490
	$(-0.2974; 0.3602)$	$(-0.7459; -0.1164)$	$(-0.8023; -0.1372)$

Table 2: **Parameter estimates for the three-factor completely affine model.** The model is estimated using daily data on zero-coupon yields from 1971 to 2006. 95%-confidence intervals are given in parentheses. Parameter values are annual and in natural units.

Parameters	Essentially affine model	Completely affine model
δ_0	-0.3869	-0.2011
	$(-0.3904; -0.3846)$	$(-0.2026; -0.1999)$
δ_X	0.0128	0.0055
	$(0.0127; 0.0128)$	$(0.0054; 0.0055)$
κ	0.1089	$2.72e-07$
	$(0.0089; 0.2573)$	$(1.13e-08; 1.63e-06)$
λ_0	3.9370	-0.0507
	$(-0.3345; 9.1777)$	$(-0.3664; 0.2729)$
λ_X	-0.1150	N.A.
	$(-0.2643; -0.0158)$	

Table 3: **Parameter estimates for the two one-factor models.** The models are estimated using daily data on zero-coupon yields from 1971 to 2006. 95%-confidence intervals are given in parentheses. Parameter values are annual and in natural units.

Panel A: Portfolio weights, $\gamma = 5$				
Inv. horizon	π_{B1}	π_{B5}	π_{B10}	π_{rf}
0	17.90	−2.54	0.44	−14.80
	(5.87; 26.72)	(−7.26; 2.63)	(−1.70; 2.65)	(−21.92; −5.37)
1	14.19	−2.25	0.92	−11.87
	(3.06; 26.03)	(−11.76; 3.88)	(−1.72; 5.17)	(−19.59; −4.02)
5	13.92	−4.85	3.04	−11.12
	(2.56; 29.58)	(−16.85; 3.12)	(−0.67; 8.22)	(−20.08, −3.76)
10	14.03	−5.85	4.23	−11.41
	(2.40; 29.61)	(−17.36; 2.67)	(0.03; 9.11)	(−20.50; −3.94)
Panel B: Portfolio weights, $\gamma = 10$				
Inv. horizon	π_{B1}	π_{B5}	π_{B10}	π_{rf}
0	8.95	−1.27	0.22	−6.90
	(2.94; 13.36)	(−3.63, 1.32)	(−0.85, 1.32)	(−10.46, −2.18)
1	7.36	−1.10	0.49	−5.73
	(1.61; 13.77)	(−6.52; 2.18)	(−0.93; 2.86)	(−9.62; −1.77)
5	7.21	−2.59	1.88	−5.51
	(0.91; 16.40)	(−10.16; 2.20)	(−0.43; 5.33)	(−10.46; −1.63)
10	7.44	−3.88	3.15	−5.71
	(0.83; 16.48)	(−11.37; 1.40)	(0.42; 6.50)	(−10.71; −1.80)

Table 4: **Optimal portfolio weights.** Panel A shows the optimal portfolios for different horizons for an investor with a risk aversion of $\gamma = 5$. Panel B shows the optimal portfolios for different horizons for an investor with a risk aversion of $\gamma = 10$. The investor is able to invest in the 1-year bond, 5-year bond, 10-year bond, and the risk-free asset. 95%-confidence intervals are given in parentheses. The portfolio weights are in natural units.

Panel A: Portfolio weights, $\gamma = 5$				
Inv. horizon	π_{B1}	π_{B5}	π_{B10}	π_{rf}
0	18.01	-3.52	0.83	-14.33
	$(5.56; 25.68)$	$(-6.91; 2.76)$	$(-1.82; 2.48)$	$(-21.01; -5.24)$
1	18.81	-3.52	0.83	-15.13
	$(6.36; 24.48)$	$(-6.91; 2.76)$	$(-1.82; 2.48)$	$(-21.81; -6.04)$
5	18.01	-2.72	0.83	-15.13
	$(5.56; 25.68)$	$(-6.11; 3.56)$	$(-1.82; 2.48)$	$(-21.81; -6.04)$
10	18.01	-3.52	1.63	-15.13
	$(5.56; 25.68)$	$(-6.91; 2.76)$	$(-1.02; 3.28)$	$(-21.81; -6.04)$
Panel B: Portfolio weights, $\gamma = 10$				
Inv. horizon	π_{B1}	π_{B5}	π_{B10}	π_{rf}
0	9.01	-1.76	0.42	-6.66
	$(2.78; 12.84)$	$(-3.45; 1.38)$	$(-0.91; 1.24)$	$(-10.01; -2.12)$
1	9.91	-1.76	0.42	-7.56
	$(3.68; 13.74)$	$(-3.45; 1.38)$	$(-0.91; 1.24)$	$(-10.91; -3.02)$
5	9.01	-0.86	0.42	-7.56
	$(2.78; 12.84)$	$(-2.55; 2.28)$	$(-0.91; 1.24)$	$(-10.91; -3.02)$
10	9.01	-1.76	1.32	-7.56
	$(2.78; 12.84)$	$(-3.45; 1.38)$	$(-0.01; 2.14)$	$(-10.91; -3.02)$

Table 5: **Suboptimal portfolio weights based on the three-factor completely affine model.** The table shows the suboptimal investment strategy for an investor who bases his investments on a three-factor completely affine term structure model. Panel A shows the suboptimal portfolio weights for different horizons for an investor with a risk aversion of $\gamma = 5$. Panel B shows the suboptimal portfolios for different horizons for an investor with a risk aversion of $\gamma = 10$. The investor is able to invest in the 1-year bond, 5-year bond, 10-year bond, and the risk-free asset. 95%-confidence intervals are given in parentheses. The portfolio weights are in natural units.

	$\gamma = 5$		$\gamma = 10$	
Inv. horizon	π_{B5}	π_{rf}	π_{B5}	π_{rf}
0	0.35	0.65	0.17	0.83
	$(-0.36; 1.67)$	$(-0.68; 1.36)$	$(-0.18; 0.84)$	$(0.16; 1.18)$
1	0.54	0.46	0.37	0.63
	$(-0.23; 1.98)$	$(-0.98; 1.22)$	$(-0.02; 1.10)$	$(-0.10; 1.01)$
5	1.31	-0.31	1.17	-0.17
	$(0.34; 3.30)$	$(-2.31; 0.66)$	$(0.66; 2.28)$	$(-1.29; 0.34)$
10	2.34	-1.24	2.17	-1.17
	$(1.06; 4.69)$	$(-3.70; -0.07)$	$(1.52; 3.65)$	$(-2.67; -0.52)$

Table 6: **Suboptimal portfolio weights based on the one-factor essentially affine model.** The table shows the suboptimal investment strategy for an investor who bases his investments on a one-factor essentially affine term structure model. The investment strategies are shown for different combinations of the investor's risk aversion, γ, and investment horizon, T. The investor is able to invest in the 5-year bond and the risk-free asset. 95%-confidence intervals are given in parentheses. The portfolio weights are in natural units.

	$\gamma = 5$		$\gamma = 10$	
Inv. horizon	π_{B5}	π_{rf}	π_{B5}	π_{rf}
0	0.37	0.63	0.19	0.81
	$(-2.01; 2.66)$	$(-1.69; 3.00)$	$(-1.00; 1.33)$	$(-0.34; 2.00)$
1	0.53	0.47	0.37	0.63
	$(-1.85; 2.82)$	$(-1.85; 2.84)$	$(-0.82; 1.51)$	$(-0.52; 1.82)$
5	1.17	-0.17	1.09	-0.09
	$(-1.21; 3.46)$	$(-2.49; 2.20)$	$(-0.10; 2.23)$	$(-1.24; 1.10)$
10	1.97	-0.97	1.99	-0.99
	$(-0.41; 4.26)$	$(-3.29; 1.40)$	$(0.80; 3.13)$	$(-2.14; 0.20)$

Table 7: **Suboptimal portfolio weights based on the one-factor completely affine model.** The table shows the suboptimal investment strategy for an investor who bases his investments on a one-factor completely affine term structure model. The investment strategies are shown for different combinations of the investor's risk aversion, γ, and investment horizon, T. The investor is able to invest in the 5-year bond and the risk-free asset. 95%-confidence intervals are given in parentheses. The portfolio weights are in natural units.

	No parameter uncertainty	Parameter uncertainty
3-factor ess.	0%	60.59%
3-factor com.	49.51%	59.49%
1-factor ess.	52.72%	54.55%
1-factor com.	48.88%	50.97%

Table 8: **The expected utility loss in a setting with and without parameter uncertainty, respectively.** All the losses are calculated under the assumption that the three-factor essentially affine model is the true model. The losses in the column named “No parameter uncertainty” are furthermore calculated under the assumption that the true set of parameter estimates are known and given by Table 1. All the losses are calculated for an investor with a risk aversion of $\gamma = 5$ and an investment horizon of 5 years.