Testing Affine Term Structure Models in case of Transaction Costs

Joost Driessen

Bertrand Melenberg

Theo Nijman

This Version: March, 2004

Testing Affine Term Structure Models in case of Transaction Costs

Abstract

We empirically analyze the impact of transaction costs on the performance of essentially affine interest rate models. We test the implied Euler restrictions and calculate the specification error bound of Hansen and Jagannathan to measure model misspecification. Using both short-maturity and long-maturity bond return data we find, under the assumption of frictionless markets, strong evidence of misspecification of affine yield models with up to three factors. Next, we incorporate transaction costs in our tests. The results show that the evidence of misspecification of essentially affine yield models disappears in case of monthly holding periods at market size transaction costs.

1 Introduction

Nowadays term structure models are used extensively for many purposes, including risk management of portfolios containing bonds and the valuation of interest-rate derivatives. Empirical tests of term structure models have therefore attracted considerable attention in the literature. In line with a large part of the empirical asset pricing literature, the tests are based on the assumption of trading in frictionless markets. In particular, the large literature on affine term structure models[1] tests these models using data on Treasury bills and bonds under the assumption of trading in frictionless markets. However, market frictions such as transaction costs are an important fact of life for investors. The implicit assumption when ignoring transaction costs is that these costs are sufficiently small, so that they do not seriously affect the empirical results. In this paper we explicitly take transaction costs into account in the empirical testing of affine term structure models, and show that including market size transaction costs can considerably affect the results of the tests.

Our approach is to test whether the stochastic discount factor of a given term structure model satisfies the Euler restrictions. These Euler restrictions are implied by the no-arbitrage assumption, and can be derived in both frictionless markets and markets with frictions. Based on these Euler restrictions, we use two approaches to analyze and test the models. First, we use Wald-type tests to test the Euler restrictions. For the frictionless case, the analysis of Euler restrictions using Wald tests is extensively discussed by Cochrane (1996, 2001). In case of transaction costs, we use tests of inequality restrictions adopting the approach developed by Kodde and Palm (1986). A disadvantage of this approach is that, if one rejects a model, there is no clear indication of the direction of misspecification, for example, which individual

---

[1]For example, Stambaugh (1988), Chen and Scott (1993), Gibbons and Ramaswamy (1993), Backus and Zin (1994), Brown and Schaefer (1994), Pearson and Sun (1994), Babbs and Nowman (1999), De Jong (2000), Backus et al. (2001a, b), Dai and Singleton (2001), and Duffee (2002). In addition, there is by now a large literature that studies models outside the affine class, including Boudhouk, Richardson, Stanton, and Whitelaw (1998), Bansal and Zhou (2001), Ahn, Dittmar, and Gallant (2002), Ahn, Dittmar, Gallant, and Gao (2002), Duarte (2003), and Leippold and Wu (2003). Dai and Singleton (2003) provide an extensive survey of this literature.

assets are possibly mispriced by the model and which are not. Also, the Wald test does not allow for a comparison of the degree of misspecification of two non-nested models that are both rejected. To overcome these problems we also consider the *specification error bound* (SEB) developed by Hansen, Heaton and Luttmer (1995) and Hansen and Jagannathan (1997). This bound measures the extent to which a model misprices a given set of assets. Hansen and Jagannathan (1997) show that this bound can be interpreted as the maximum pricing error for all portfolios that can be constructed from the assets under consideration. This specification error bound allows for direct comparison across (non-nested) models and the method indicates which (portfolios of) assets contribute most to the misspecification. Hansen, Heaton and Luttmer (1995) extend the setup of Hansen and Jagannathan (1997) to allow for market frictions. We apply their approach to affine term structure models and compare the results with standard tests using the Euler restrictions.

Our work is related to Luttmer (1996) and He and Modest (1995), who both analyze the influence of transaction costs and other market frictions on the size of the *volatility bounds* of Hansen and Jagannathan (1991), that give a lower bound on the variability of valid stochastic discount factors. Luttmer (1996) finds that small transaction costs greatly influence the size of the volatility bounds; especially the volatility bounds based on T-bill returns are very sensitive to the size of transaction costs. The results of Luttmer (1996) imply that the conclusion of rejection of several asset pricing models in Hansen and Jagannathan (1991), based on the volatility bounds, changes if transaction costs are taken into account. Our work extends the work of Luttmer (1996), because the volatility bound is a special case of the specification error bound, and because Luttmer (1996) focuses on consumption-based asset pricing models, whereas we analyze bond pricing models and bond returns.

The bond pricing models that we consider are discrete-time versions of the affine yield models of Duffie and Kan (1996) and the extension to essentially affine models due to Duffee (2002). These latter models deviate from the Duffie and Kan (1996) models – referred to as completely affine yield models in this context– by allowing the market price of risk to depend in a non-affine way on the factors. In our investigation we follow the classification of these models as proposed by Dai and Singleton (2000),

consisting of a partitioning into subclasses –for each given number of factors– depending on the way the factors affect the volatility of the process generating the uncertainty. We consider two and three factor models, since these are commonly studied in the literature mentioned above. Given the number of factors, we compare the various subclasses, and select a most preferred subclass, for which we present our results.

To relate our tests of Euler restrictions to existing empirical work on affine models, we show that the Euler restrictions can be rewritten into restrictions on unconditional and conditional expected returns on bonds of different maturities. In line with the literature on tests of the expectation hypothesis (Fama and Bliss (1987) and Campbell and Shiller (1991)), which shows that the spread between a long-term and short-term interest rate predicts future bond returns, we choose as conditioning variable this yield spread. This way, our test restrictions on conditional expected bond returns are related to Dai and Singleton (2001) and Duffee (2002), who analyze to what extent essentially affine models can reproduce the predictability of bond returns by the yield spread.[2] However, compared to these two articles, our set of test restrictions is larger since we also include restrictions on the term structure of unconditional expected bond returns in our tests.

Before discussing the empirical results, we perform a simulation analysis to analyze the power of the Wald-test on the Euler restrictions. We simulate a three-factor affine model, adding transaction costs, and then test a two-factor affine model (both for the case with frictionless markets and for the case with transaction costs). Although the power of the tests weakens somewhat if the test incorporates transaction costs, we find that the test does have reasonable power to reject the two-factor model in this setup.

Our empirical results indicate that, assuming no market frictions, the term structure of average bond returns is less smooth than predicted by the affine models: both the completely affine yield models and the essentially affine yield models significantly misprice the returns on portfolios that contain both extreme long and short positions in near-maturity bonds. In particular, we find that two-factor models

---

[2]Ahn, Dittmar, and Gallant (2002), Duarte (2003), and Leippold and Wu (2003) analyze the predictability of bond returns (or yield changes) from the perspective of models outside the affine class of term structure models.

are clearly rejected, in line with the existing literature. Furthermore, even the more general essentially affine three-factor models are statistically rejected if we test the Euler restrictions under the assumption of frictionless markets. Dai and Singleton (2001) and Duffee (2002) show that some essentially-affine models are capable of fitting the observed predictability of bond returns. Our results show that, if both restrictions on unconditional and conditional expected returns on bonds with different maturities are tested, all essentially affine models are rejected.

Instead of extending the essentially affine three-factor model further, we allow for transaction costs in our tests. We find that, when transaction costs are of market size, the conclusions above need a more carefully balanced appraisal. In case of a monthly holding period, the evidence of misspecification of the considered models disappears when these transaction costs are included. Because of the transaction costs, the portfolios with both long and short positions in T-bills and bonds are no longer mispriced. For quarterly holding periods and market size transaction costs, the results are mixed: the models are not rejected on the basis of data on long-maturity bond returns, but these models do misprice the short-maturity T-bills. However, Duffee (1996) provides evidence that T-bill returns with very short maturities contain a large idiosyncratic component, possibly due to market segmentation. This might partially explain the difficulty the models have in pricing short-maturity T-bills.

We supplement our empirical analysis with a second simulation study in order to gain further insights into our findings. We simulate a two-factor affine model with transaction costs, and test three-factor affine models without allowing for transaction costs. The results show that the three-factor models are clearly rejected in this case. This is in line with the empirical results where we also find that three-factor models are rejected in frictionless markets. These results support our conclusion that the presence of transaction costs can lead to a rejection of appropriate models if one tests under the assumption of frictionless markets.

The remainder of this paper is organized as follows. In section 2, we briefly review the literature on affine term structure models. In section 3, we first summarize the literature on asset pricing in markets with frictions, then we describe a Wald-test of the Euler restrictions in such a market with frictions, and

we discuss the specification error bound. In section 4, after describing our dataset and estimation procedures, we present the empirical test results and results from the simulation studies. In section 5 we summarize and conclude.

2 Affine Term Structure Models

Let the n-dimensional vector $R_{t+1} = (R_{1,t+1},...,R_{n,t+1})'$ contain the gross returns from time t to time $t+1$ of n assets (in our case bonds of n different maturities). In the empirical analysis, we analyze both monthly and quarterly holding periods, so that the returns R_{t+1} are either monthly returns or quarterly returns. Also, let Y_{t+1} denote a stochastic discount factor (SDF), such that, in case of no arbitrage opportunities in terms of the n assets and no market frictions,

$$E_t[R_{i,t+1} \, Y_{t+1}] \; = \; 1, \quad i=1,..,n, \tag{1}$$

with Y_{t+1} strictly positive, and where the expectation is conditional on the information set at time t (see, for example, Campbell, Lo and MacKinlay (1997), Chapter 11).

Duffie and Kan (DK, 1996) describe the class of continuous-time multi-factor term structure models, that imply an affine relationship between interest rates and a vector of state variables. As reported by Duffee (2002), these 'completely' affine models might produce poor forecasts of future Treasury yields, attributed to the fact that the risk compensation cannot vary independently of the interest rate volatility. For this reason Duffee (2002) proposes the class of 'essentially affine' models, as extension of the 'completely' affine models of Duffie and Kan (1996). In both the completely and essentially affine models the conditional means and covariances of the factors are affine functions of the current factor values. However, to avoid the drawback of the completely affine models, the market prices of risk are not affine anymore for essentially affine models.

Our setup is in discrete time. We will use discrete-time versions of these models, as described by Backus et al. (2001a, b), see also Campbell, Lo and MacKinley (1997). Although various discrete-time versions of the continuous-time models are possible, the one proposed by Backus et al. (2000a, b) seems to be the most natural one. This approach assumes a conditionally normal distribution for an N-dimensional vector of state variables a_{t+1}. We extend this discrete-time setup to include the essentially affine specification, by transforming the continuous-time model of Duffee to our discrete-time framework. Denoting the log-SDF, $y_{t+1} = \log(Y_{t+1})$, the N-factor essentially affine discrete-time model can be written as[3]

$$-y_{t+1} = \varphi_0 + \varphi_1{}' a_t + C(a_t) + \gamma{}' \xi_{t+1} + a_t{}' \Gamma' V(a_t)^{-1} \xi_{t+1}, \tag{2}$$

with

$$
\begin{aligned}
C(a_t) &= a_t{}' \Gamma' V(a_t)^{-1} \Gamma a_t / 2 \\
a_{t+1} &= \mu + \Lambda (a_t - \mu) + \Sigma \, \xi_{t+1} \\
\xi_{t+1} \mid I_t &\sim N(0, V(a_t)) \\
V(a_t) &= diag(\alpha_1 + \beta_1{}' a_t, \ldots, \alpha_N + \beta_N{}' a_t).
\end{aligned}
\tag{3}
$$

Here ξ_{t+1} represents an N-dimensional conditionally normally distributed random vector with zero conditional mean and conditional variance matrix $v(a_t)$, Λ, Γ, and Σ are $N \times N$-matrices containing unknown parameters, φ_0 is a one-dimensional parameter, φ_1, $\alpha = (\alpha_1, \ldots, \alpha_N)'$, $\beta_1, \ldots, \beta_N$, μ, and γ are N-dimensional unknown parameter vectors, and I_t represents the information set of time t. If $\Gamma = 0$, equation (2) reduces to the completely affine models of DK. In this case, the term $\gamma{}' \xi_{t+1}$ captures the market prices

---

[3]Notice a small difference with Campbell, Lo and MacKinley (1997), whose specification of the log-SDF also contains a normally distributed variable that is independent from ξ. This variable only influences the mean of the yield curve in a way that is very similar to the way the mean of the state-variable influences the mean yield curve. Following Backus et al. (2000a, b) we do not include this variable in our analysis (also in line with Backus and Zin (1994), and Bansal (1997)), allowing us in a straightforward way to calculate the SDF in terms of observables.

of risk, as it measures the sensitivity of the SDF (and, thus, bond returns) for the underlying factors.[4] If $\Gamma \neq 0$, the term $\gamma'\xi_{t+1} + a_t'\Gamma'V(a_t)^{-1}\xi_{t+1}$ captures the market price of risk[5].

In the application we consider two and three factor models. We follow the classification according to Dai and Singleton (2000), indicated by $A_m(N)$, where N denotes the number of factors, and m the number of factors affecting the volatility. For instance, in case of two-factor completely affine models, this means that we can distinguish between the following three models

- $A_0(2)$ model: Factors with constant volatility, i.e., a Vasicek (1977)-type model.

- $A_1(2)$ model: One factor with constant volatility and one factor with square root volatility.

- $A_2(2)$ model: Two factors with square root volatility, i.e., a Cox, Ingersoll, Ross (1985)-type model.

Not all parameters in these affine models are identified. We follow Dai and Singleton (2000) to impose appropriate normalizations. We refer to the appendix A for further details.

As a comparison to the completely affine models we also analyze the essentially affine three-factor models, using the classification by Duffee (2002), $EA_0(3)$, $EA_1(3)$, and $EA_2(3)$. The $EA_3(3)$ model is the same as the $A_3(3)$ model. Here $EA_m(N)$ means an N-factor essentially affine model, where m factors affect the conditional volatility. We follow Duffee (2002) in the normalization, see appendix A for further details.

Using equations (1)-(3), one can derive that bond prices are exponential-affine functions of the state variable a_t,

$$-\log P_{nt} = n r_{nt} = A_n + B_n' a_t, \tag{4}$$

---

[4]Notice that, since ξ_{t+1} follows a normal distribution with mean zero and covariance matrix $V(a_t)$, it would make sense to define $V(a_t)^{1/2}\gamma$ as the vector containing as components the market prices of risk.

[5]The term $C(a_t)$ appears since the discrete time approximation is based on log-s.

where P_{nt} is the price of an n-period zero-coupon bond at time t, and r_{nt} is the corresponding interest rate. The factor loadings A_n and B_n are functions of the underlying parameters, and do not depend on time. This equation also holds for the essentially affine models, so that the analytical tractability is maintained.

When estimating the affine yield models, we allow for a measurement error in the yields, following, for example, the approach of Duan and Simonato (1999) or De Jong (2000), meaning that, instead of (4), we use

$$r_{nt} = \frac{1}{n}A_n + \frac{1}{n}B_n' a_t + \varepsilon_{nt}, \tag{5}$$

where ε_{nt} represents the measurement error. As discussed in the data section, we use 'smoothed' interest rates for estimation, which motivates the fact that we allow for measurement error. Based on (2), (3), and (5), together with appropriate distributional assumptions, we estimate the unknown parameters by Maximum Likelihood using the Kalman filter. The estimation procedure is described in detail in appendix B.

Besides parameter estimates, the Kalman filter provides estimates for the factor values at all dates in the data set, which can be used to obtain estimated values for the SDF at all dates, which is required for the empirical testing procedure. We rewrite the innovation in equation (3) as

$$\xi_{t+1} = \Sigma^{-1}(a_{t+1} - \Lambda(a_t - \mu)) \tag{6}$$

We then substitute this into equation (2) to obtain an estimate for the SDF. For instance, in case of a completely affine yield model, we see that the SDF is given by

$$-y_{t+1} = \varphi_0 - \gamma'\Sigma^{-1}\Lambda\mu + (\varphi_1' - \gamma'\Sigma^{-1}\Lambda)a_t + \gamma'\Sigma^{-1}a_{t+1} \tag{7}$$

A similar expression holds for the essentially affine models.

In the next section we describe how we test these affine term structure models in case of transaction costs. These tests will be based on Euler restrictions (as in equation (1) for frictionless markets). Most

empirical work on term structure models has tested these models using the price restrictions in equation (4).[6] In this section we have shown that these price restrictions are derived from the Euler restrictions, so that there is a direct link between the Euler restrictions and the price restrictions. One would therefore not expect important differences between tests based on equation (4) versus equation (1).

3 Testing the Models in case of Transaction Costs

3.1 Price implications in case of transaction cost

Without transaction costs the models can easily be tested by verifying whether moment restrictions implied by (1) are satisfied. However, transaction costs are a fact of life. With transaction costs, the moment restrictions implied by (1) are too strong, so that rejection of these moment restriction is no longer an indication of model misspecification. To see this, consider first short-selling constraints on the assets. In this case, absence of arbitrage opportunities requires the existence of a strictly positive SDF satisfying (instead of (1))

$$E_t[R_{i,t+1}\, Y_{t+1}] \le 1, \quad i=1,...,n, \tag{8}$$

see, for example, Jouini and Kallal (1995) or Luttmer (1996).

When considering transaction costs, we restrict ourselves to the case of a proportional spread s that is equal at the ask and bid side, and the same for all assets under consideration. Let P_{it} denote the midprice of asset i at time t. Then the gross return on taking a long position is equal to

$$\frac{(1-s/2)\; P_{i,t+1}}{(1+s/2)\; P_{it}} \;\equiv\; \tau^l \frac{P_{i,t+1}}{P_{it}} \;\equiv\; \tau^l R_{i,t+1}, \tag{9}$$

---

[6]An important exception is Gibbons and Ramaswamy (1993), who also test term structure models using Euler equations.

and for short positions the gross return is equal to

$$\frac{(1+s/2)\ P_{i,t+1}}{(1-s/2)\ P_{it}} \equiv \tau^s\ \frac{P_{i,t+1}}{P_{it}} \equiv \tau^s R_{i,t+1}. \tag{10}$$

In testing, transaction costs can be taken into account by rewriting the problem as one with restrictions on short and long positions (see Luttmer (1996)) and introducing separate assets for a long position in asset i with return $\tau^l R_{i,t+1}$ and for a short position with return $\tau^s R_{i,t+1}$. As a consequence, the absence of arbitrage opportunities in the presence of transaction costs requires the existence of a strictly positive SDF Y_{t+1} such that

$$\frac{1}{\tau^s} \leq E_t[Y_{t+1} R_{i,t+1}] \leq \frac{1}{\tau^l},\ i=1,...,n. \tag{11}$$

In the empirical analysis we use unconditional Euler restrictions. Following Hansen and Jagannathan (1991) and many others, we incorporate conditional information by constructing returns on managed portfolios with payoffs $x_{t+1} = z_t \otimes R_{t+1}$ and corresponding price vector $q_t = z_t \otimes \iota_n$, where z_t is an m-dimensional vector with variables that are in the information set at time t.[7] The implied unconditional Euler restrictions are

$$\frac{1}{\tau^s} \leq \frac{E[Y_{t+1} x_{i,t+1}]}{E[q_{it}]} \leq \frac{1}{\tau^l},\ i=1,...,m \times n, \tag{12}$$

where $x_{i,t+1}$ and q_{it} represent the ith component of x_{t+1} and q_t, respectively. In the sequel, we shall refer to the vector x_{t+1} as the vector of returns, and we shall denote the number of returns in (12) simply by n, instead of $m \times n$, to avoid too cumbersome notation.

In general, this 'multiplicative' approach may not be an optimal way of incorporating conditional information. Indeed, for the volatility bounds of Hansen and Jagannathan (1991), Ferson and Siegel

---

[7]In the empirical analysis section we present our choice of z_t.

(2003) discuss how to use conditional information optimally. However, the question of how to incorporate conditional information in an optimal way in case of the specification error bound (to be discussed in Subsection 3.3) does not seem to be resolved yet. Because of this, and because of the simplicity of the multiplicative approach, we shall use this approach.

3.2 Testing using a Wald-Test

For every affine term structure model, Wald-type tests of the Euler restrictions for both the frictionless case (implied by equation (1)) and the case with transaction costs (following from equation (11)) are relatively straightforward to implement. For the case of transaction costs, the inequality constraints can be tested along the lines of Kodde and Palm (1986). In this section, we briefly consider how such a test can be performed. The test for frictionless markets is a special case. We start by assuming that the SDF Y_{t+1} is fully observed.

The implied null hypothesis we test is that the SDF satisfies the Euler restrictions (12). Given T time-series observations on the n-dimensional vector of returns x_{t+1} and a candidate SDF, we estimate the ratio of expectations in (12) by its sample analogue

$$\hat{v}_i \equiv \frac{\frac{1}{T}\sum_{t=1}^{T} Y_{t+1} x_{i,t+1}}{\frac{1}{T}\sum_{t=1}^{T} q_{it}}, \quad i=1,...,n. \tag{13}$$

Then we use as test-statistic ξ_w given by

$$\xi_w = \min_{v \in \mathbb{R}^n} T(\hat{v}-v)' \hat{W}^{-1} (\hat{v}-v)$$

$$s.t. \quad \frac{1}{\tau^s} \leq v_i \leq \frac{1}{\tau^l}, \quad i=1,...,n, \tag{14}$$

where $v = (v_1,...,v_n)'$, and where $\hat{W}$ denotes a consistent estimator of the asymptotic covariance matrix of $\hat{v} = (\hat{v}_1,...,\hat{v}_n)'$. As suggested by Wolak (1989, 1991), we interpret this test as a local test of the

inequality constraints, meaning that -from an asymptotic point of view- for each $i=1,...,n$, at most one of the inequalities in (14) will be relevant.[8] As discussed in appendix C, the test is then a straightforward special case of the test proposed by Kodde and Palm (1986).

In the absence of transaction costs, the test-statistic ξ_w reduces to the J-statistic of Hansen (1982), and follows, under the null hypothesis, asymptotically a chi-square distribution with n degrees of freedom. In case of transaction costs, it follows from Kodde and Palm (1986) that, under the null hypothesis, this test-statistic is asymptotically distributed as a mixture of chi-square distributions. In this case simulation can be used to obtain p-values for a given value of the test-statistic.

In the empirical application the SDF Y_{t+1} contains unknown parameters, which have to be estimated. Estimation of the SDF means that the limiting distribution under the null hypothesis of the test statistic discussed above has to be adapted. We refer to appendix C for further details.

A disadvantage of the testing methodology of this subsection is that, if a model is rejected, there is little indication of the direction of the misspecification. Also, if one rejects two non-nested models, no indication is obtained whether one model is more misspecified than the other. In the next subsection, we will argue that the use of the specification error bound overcomes these problems.

3.3 Testing using the Specification Error Bound

As stressed by Hansen and Jagannathan (1997), an asset pricing model is an approximation of reality and, therefore, it will typically not exactly satisfy the Euler restrictions in an empirical analysis. These authors propose to measure the size of misspecification of a given proxy model, with SDF Y_{t+1}, by measuring in some way the pricing errors of this proxy model. In this section, we briefly describe the part of their approach that is relevant for our application. Again, we start by assuming that the SDF Y_{t+1} is fully observable.

---

[8] A global interpretation of our test procedure would imply that we overestimate the size of transaction costs that is needed to avoid statistical rejection of the model, or equivalently, that we would underestimate the influence of transaction costs on model misspecification.

In our case, the proxy model is given by one of the models that we described in section 2. We start by introducing the set M of admissible SDFs consisting of random variables m_{t+1} (which are in the information set at time $t+1$) that satisfy the Euler restrictions

$$\frac{E[q_{it}]}{\tau^s} \leq E[m_{t+1} x_{i,t+1}] \leq \frac{E[q_{it}]}{\tau^l}, \quad i=1,...,n. \tag{15}$$

Thus, an SDF is admissible if it prices all (linear combinations of) assets under consideration correctly. The SDF Y_{t+1} that is associated with the proposed model can be used to calculate model prices of the payoffs, that, in general, may not satisfy the restrictions in (15). Hansen and Jagannathan (1997) propose to measure the size of this misspecification by

$$\delta^2 = \min_{m_{t+1} \in M} E[(Y_{t+1} - m_{t+1})^2]. \tag{16}$$

The square root of (16) is called the Specification Error Bound (SEB), and can be interpreted as a (minimum) distance between the proxy SDF Y_{t+1} and the set of admissible SDFs.[9]

For the case without market frictions (i.e., $\tau^s = \tau^l = 1$ in (11)), Hansen and Jagannathan (1997) show that the SEB following from equation (16) has an interpretation as the maximal pricing error of all portfolios in the n assets

$$\delta = \max_{\lambda \in \mathbb{R}^n} \quad |E[Y_{t+1}(\lambda^{/} x_{t+1}) - \lambda^{/} q_t]|$$
$$s.t. \quad E[(\lambda^{/} x_{t+1})^2] = 1 \tag{17}$$

It is easy to show that this interpretation of the SEB still holds in the case of transaction costs. More precisely, given the set M defined by (15), one can show that, with market frictions of the form (11), δ

---

[9]Hansen and Jagannathan (1997) also introduce a bound where the set of admissible SDFs only contains SDFs with the same unconditional mean as the proxy SDF, and show that this condition is automatically satisfied if one analyzes models with a stochastic discount factor that contains an additive, unknown constant term, that is chosen such as to minimize the SEB. We do not analyze stochastic discount factors with this property, and we also do not impose this restriction on the mean of the SDF, because this would imply that any model that we analyze prices the return of a one-period bond without error.

satisfies

$$\delta = \max_{\lambda \in \mathbb{R}^n} \min_{v \in \mathbb{R}^n} |E[Y_{t+1}(\lambda' x_{t+1}) - \lambda' v]|$$

$$s.t. \quad E[(\lambda' x_{t+1})^2] = 1$$
$$\frac{E[q_{it}]}{\tau^s} \leq v_i \leq \frac{E[q_{it}]}{\tau^l}, \quad i = 1, ..., n. \tag{18}$$

Both (17) and (18) show that δ gives a bound on pricing errors of portfolio payoffs that are normalized in a particular way. Note that this normalization does not imply that the components or 'weights' in λ, which are equal to the Kuhn-Tucker multipliers of the binding Euler restrictions (see Hansen and Jagannathan (1997)), sum up to one.

A slight modification of a frictionless result in Hansen and Jagannathan (1997), adapted to the case of transaction costs, reveals that the SEB of (18) can also be calculated as

$$\delta^2 = \min_{v \in \mathbb{R}^n} \quad E[Y_{t+1} x_{t+1} - v]' E[x_{t+1} x_{t+1}']^{-1} E[Y_{t+1} x_{t+1} - v]$$

$$s.t. \quad \frac{E[q_{it}]}{\tau^s} \leq v_i \leq \frac{E[q_{it}]}{\tau^l}, \quad i = 1, ..., n \tag{19}$$

Comparing this with equation (14) shows that the SEB is closely related to the population analogue of the Wald test-statistic. The only difference is the weighting matrix.

By replacing population moments with their sample analogues in equation (19), an estimate $\hat{\delta}$ for the SEB can be obtained. Hansen, Heaton and Luttmer (1995) show, under the assumption that the true δ is strictly positive, that this estimator has asymptotically a normal limiting distribution; they also provide a consistent estimate for the asymptotic variance[10]. The assumption that the true bound is strictly positive is crucially different from the setup of the Wald-test, where the null hypothesis is that the model is correctly specified.

---

[10] If the true δ is equal to zero, Hansen, Heaton and Luttmer (1995) argue that the limit distribution is mixed chi-square if there are no transaction costs. This test is then less efficient than the Wald-test discussed in the previous subsection.

Thus, although the mathematical difference between the Wald test-statistic and the SEB is only the form of the weighting matrix, the Wald-test and the SEB are two complementary approaches. The Wald-test allows for efficient statistical testing based on the Euler restrictions of a given model, but it does not provide information on the direction of misspecification. If the model is misspecified, the properties of the tests are not easy to derive. For the SEB, it is *a priori* accepted that the model is misspecified; therefore, the size of misspecification is measured, along with the contributions of individual assets to this misspecification size by means of the Kuhn-Tucker Multipliers.

In the empirical application, we do not observe the SDF Y_{t+1}, but, instead, we have to estimate it. The preliminary round of estimation requires that the limit distribution of the SEB has to be adapted in a similar way as the Wald test.

4 Empirical Results

4.1 Data

The dataset that we use contains monthly data on interest rates and bond holding returns. The interest rate data are drawn from the CRSP Fama Files, and consist of interest rates of maturities ranging from 1 month to 5 years. The short-maturity interest rates are derived from T-bill prices, and the long-maturity interest rates are calculated from bond prices. We use a subsample from 1972-1997[11], consisting of 312 monthly observations. In table 1 some basic sample statistics of the data are presented. These interest rate data are used for the first-step Kalman filter estimation discussed below.

< Insert table 1 around here >

The monthly holding returns data that we use also come from the CRSP Fama Files. For maturities up to one year, we use the nominal holding returns that are calculated from T-bill prices. For longer maturities, we use the returns on the so-called maturity portfolios available in the CRSP Fama Files,

---

[11]Before 1972 there are missing observations for some variables in the data.

which are constructed from bonds whose maturity lies in a given interval. The intervals we use are: 2 to 3 years, 4 to 5 years, 5 to 10 years, and larger than 10 years. These returns exactly represent returns on tradable portfolios. Again we use the subsample from 1972-1997. In table 2 we report some sample properties of these data. From this table, it is clear that the average holding returns differ considerably for the various short maturities, whereas the differences in average holding returns for the long-maturity assets are quite small, relative to the standard deviations and the difference in maturity.

< Insert table 2 around here >

In table 3 we report information on the bid-ask spreads on T-bill prices, which are derived from the CRSP data. It follows that the size of the transaction costs due to the bid-ask spread is around 1.5 basis points, averaged over time and over all T-bills. Table 3 also shows that the bid-ask spreads have decreased considerably during the last 25 years. For bonds, we refer to Chakravarty and Sarkar (1999), who report a bid-ask spread of government bonds (with maturities ranging from 10 months to 30 years) of around 11 cents when reported on the basis of a $100 par value.

< Insert table 3 around here >

The following sets of assets returns will be used in the empirical analysis:

1. *Short-Maturities Asset Set:* Four T-bills with maturities of 1, 3, 6, and 9 months.

2. *Long-Maturities Asset Set:* Four bond portfolios with maturity intervals equal to 2 to 3 years, 4 to 5 years, 5 to 10 years, and larger than 10 years.

3. *All-Maturities Asset Set:* Set 1 and set 2.

Thus, we consider three subsets of assets, one that contains only short-maturity T-bills, another one that contains only long-maturity bonds and a third one that contains bonds of both short and long maturities. The maturities of the T-bills are the same as in Luttmer (1996). As mentioned earlier, we will both use

monthly and quarterly returns on these assets to perform the Wald-tests and calculate the SEBs.

4.2 First round Maximum Likelihood estimation results

Before being able to test the models, we first need to estimate the SDFs $Y_{t+1} = Y_{t+1}(\theta)$ of the various models, see equation (7) for the completely affine yield models. We use Maximum Likelihood based on the Kalman filter to estimate the parameter vector θ. This estimation procedure is described in detail in appendix B. Some of the affine subfamilies contain many parameters (more than 20 in the three-factor case). Therefore, following Duffee (2002), we first estimate a completely unrestricted specification for each subfamily, and set parameters with t-ratios smaller than 1 in absolute size equal to zero in a second estimation step, except parameters that are restricted to be strictly positive (the diagonal components of $(I-\Lambda)$, and the unconditional means of square-root factors, see equations (2) and (3)). To prevent an overload of tables, we only report results for the most preferred models for three sets of models (two- and three-factor completely affine models, and three-factor essentially affine models), where we use as criterion to select the most preferred models the monthly SEB for the all-maturities asset set. The estimation results for these 'most preferred' models are presented in Table 4.[12]

< Insert table 4 around here >

For both the two- and three-factor models, one factor has very slow mean reversion, so that this factor causes almost parallel term structure movements. The other factors revert much faster to their long-term means, and therefore mainly influence short-maturity interest rates (in line with results of De Jong (2000) and Duffee (2002)). The model fit seems very reasonable, with the best performance for the $EA_1(3)$ model, which has a mean absolute yield error of 6 basis points across all maturities.

4.3 Conditional information

To perform the tests described in section 3, we also need to specify the conditional information. Recall

---

[12]The other estimation results are available upon request.

that $x_{t+1} = z_t \otimes R_{t+1}$ and $q_t = z_t \otimes \iota_n$, where z_t is a vector with variables that are in the information set at time t. Following Luttmer (1996), we construct these conditioning variables in such a way that they are always positive, so that short-selling constraints or transaction costs are straightforward to impose on the 'conditional' assets as well. Given the empirical evidence that the yield spread predicts future interest rate movements (Fama and Bliss (1987) and Campbell and Shiller (1991)), we choose to use the yield spread as conditioning variable. More precisely, in case of the *Short-Maturities Asset Set* the conditional information consists of a constant and the ratio of the 1-year and the 3-month interest rate ('the short yield spread'); in case of the *Long-Maturities Asset Set* the conditional information consists of a constant and the ratio of the 5-year interest rate and the 1-year interest rate ('the long yield spread'); and in case of the *All-Maturities Asset Set* the conditional information consists of a constant, the short yield spread for T-bills and the long yield spread for the maturity bond portfolios. This implies that the *Short-Maturities Asset Set* and the *Long-Maturities Asset Set* both contain 8 returns, whereas the *All-Maturities Asset Set* contains 16 returns.

Dai and Singleton (2001) and Duffee (2002) analyze to what extent essentially affine models can fit the observed predictability of bond returns by the yield spread. Effectively, they assess the difference between the empirical covariance and the model-implied covariance of the current yield spread with the subsequent bond return. To relate these covariance moment restrictions to the Euler equations, we rewrite the Euler restrictions in equation (12) as follows

$$\frac{E[q_{it}]}{\tau^s E[Y_{t+1}]} - \frac{Cov[Y_{t+1}, x_{i,t+1}]}{E[Y_{t+1}]} \leq E[x_{i,t+1}] \leq \frac{E[q_{it}]}{\tau^l E[Y_{t+1}]} - \frac{Cov[Y_{t+1}, x_{i,t+1}]}{E[Y_{t+1}]}, \quad i=1,...,m \times n. \tag{20}$$

This equation shows that the Euler equations can be rewritten into restrictions on unconditional expected bond returns (since z_t includes a constant), and on conditional expected bond returns (since z_t includes the yield spread). We thus test whether the average (conditional) bond returns satisfy the lower and upper bounds in (20), which are implied by a specific pricing kernel. Given that we include the unconditional expected bond returns as moment restrictions, the restrictions on the conditional expected bond returns

are directly related to restrictions on the covariance between the yield spread and bond returns, as analyzed by Dai and Singleton (2001) and Duffee (2002). Compared to these two articles, our set of test restrictions is thus larger, since we also include the term structure of unconditional expected bond returns.

Given the estimated SDFs of the various models, we can get a first impression of the model accuracy by calculating the pricing errors of the managed portfolios constructed from the *All-Maturities Asset Set* for each of the models. The pricing error for the managed portfolio based on asset i is defined as

$$\frac{1}{T}\sum_{t=1}^{T} Y_{t+1}\, x_{i,t+1} \; - \; \frac{1}{T}\sum_{t=1}^{T} q_{it} \tag{21}$$

In table 5 we present the average pricing errors together with corresponding t-values, calculated over all managed portfolios on the basis of the *All-Maturities Asset Set* for the different models, in case of a monthly holding period. The most preferred two- and three-factor completely affine yield models are performing more or less equally, while the three-factor most preferred essentially affine yield model seems to perform slightly better. Table 5 also contains the pricing error correlation across the models to show which models are close and which are more apart. As can be seen from this table, there is a substantial correlation between the pricing errors of the various preferred models.

< Insert table 5 around here >

4.4 Power of Wald test

Before discussing the empirical results, we use simulation to analyze the power of the Wald test methodology. The simulation setup consists of the following steps. First, we simulate a three-factor model and add transaction costs of 1 basis point to the simulated bond prices in a random way across time and maturities (i.e., the bond price is multiplied with $1+\tau$ or $1-\tau$ with equal probability). From these bond prices, we then construct monthly time series of interest rates and bond returns (for a 25 years time period). We then use the simulated data to estimate and test both the (correct) three-factor model to assess the size of the test, and a two-factor model to assess the power of the test. The two-factor model is a

-19-

special case of the three-factor model. This procedure is repeated 1000 times.

< Insert table 6 around here >

We analyze three subfamilies of affine three-factor models. Panel A of Table 6 gives the results for the size of the test. In case of tests that assume frictionless markets, the simulation results show that the tests tend to reject models too often (given a confidence level of 95% and the 25-year sample size). Once we incorporate the assumed transaction cost of 1 basis point in our Wald test, the empirical size of the test is closer to the 5% level, although in this case there are too few rejections of the model. Next, we turn to Panel B of Table 6 which shows the power of the Wald test (and the associated SEB). For the tests of the two-factor models that assume frictionless markets, we find a clear rejection of the two-factor models in almost all of the 100 simulations. The SEB for the two-factor models is also large and comparable in size to the SEBs that we find empirically for the two-factor models in case of frictionless markets (as discussed in subsection 4.4). We also test the two-factor models allowing for the presence of transaction costs of 1 basis point. In this case, we find that the rejection rate is somewhat smaller compared to the frictionless market test. However, the Wald test still rejects two-factor models in the majority of the simulations, and the SEBs of the two-factor models are large. Finally, Table 6 shows that a 95% rejection rate is obtained if one assumes transaction costs of about 0.1 basis point. In sum, these simulation results show that the Wald test has reasonable power to distinguish two-factor models from three-factor models for the sample size used in this paper, even if we allow for the presence of transaction costs.

4.5 Empirical test results

In this subsection we present empirical results for the specification tests, first of all, for a setup without transaction costs, and then with transaction costs, and both for monthly and quarterly holding periods.

We start with the case *without transaction costs*.

< Insert table 7 around here >

-20-

In table 7, we present the results of the Wald-test.[13] As the table shows, the Wald-test (conducted at the 5% level) on the frictionless Euler restrictions rejects all (preferred) models for all asset sets and for both monthly and quarterly holding periods with only one exception: the most preferred completely affine three-factor yield model applied to the long-maturities case, using a monthly holding period. This model is however strongly rejected on the basis of the full set of assets. Thus, when confronted with the frictionless Euler restrictions, we have to conclude that generally speaking the models seem to be misspecified. For the two-factor models, this is in line with other research: using different test procedures and different data, Dai and Singleton (2000) and De Jong (2000) also reject two-factor term structure models. Interestingly, even the three-factor essentially affine model is rejected statistically. Duffee (2002) has shown that the fit of this model is clearly better than the completely affine models, but that the essentially affine models still have problems with fitting both the time-variation in conditional variances and interest rate risk premia. Our test results complement Duffee's results since we show that there is also statistical evidence that the essentially affine models are misspecified, if we assume frictionless markets and focus on the term structure of expected (conditional) bond returns.

< Insert table 8 around here >

Table 8 reports the SEBs of the various preferred models. As can be seen from this table, the SEBs are large and far from zero. Notice that the difference between the various models is quite small, in line with the correlation results reported in table 5.

The SEBs in case of the T-bills are much larger than the bounds based on long-maturity bonds. As Luttmer (1996) notices, an explanation for the high T-bill bounds is that, because the holding returns on the different T-bills are highly correlated, differences in average holding returns on these T-bills can lead to something close to an arbitrage opportunity. Thus, the admissible set of SDFs is relatively small. For the long-maturity bonds the differences in average holding returns are not very large, especially relative

---

[13]For quarterly holding periods we use the Newey-West (1987) method to estimate W, with two lags, in order to correct for the overlapping nature of the quarterly pricing errors. We also estimated the matrix W using Newey-West (1987) with ten lags. This hardly changes the results.

to the volatility of the holding returns, and thus the admissible set of SDFs is larger in this case.

The economic significance of the estimated bounds under the assumption of frictionless markets is large. For example, based on the SEB results for the All-Maturities Asset Set and the preferred three-factor essentially affine model, we can conclude that for this model there exists a portfolio, normalized as in equation (17), with a pricing error of 0.617. This portfolio has an observed (mid)price of 0.646, whereas the model assigns a price of 0.029 to this portfolio. The multipliers in equation (17) define the weights of this maximum pricing error portfolio and show that this most severely mispriced portfolio is characterized by extreme short and long positions in bonds with adjacent maturities (not reported). This implies that the model is rejected in this frictionless setting because the observed behaviour of bond returns of different maturities is less smooth than implied by the model. In his study of Euler equations for equity returns, Cochrane (1996) also finds that portfolios with long and short equity positions are largely mispriced.

To obtain further insight in these results, we calculate the pricing errors for two types of portfolios: portfolios in only one T-bill or bond, and two-asset portfolios that have a long position in one T-bill (bond) and an equally large short position in another T-bill (bond). To facilitate the comparison between these portfolio pricing errors and the SEBs in table 8, we normalize these portfolio weights in the same way as the SEB-weights λ in equation (18) are normalized. Table 9 presents the monthly pricing errors, in case of the preferred essentially affine three-factor model. It follows that individual T-bill and bond returns have low pricing errors; the normalized pricing errors are much smaller than 0.01 for all assets. The normalized pricing errors for the portfolios in two assets are much larger than the pricing errors for the individual assets, especially for the T-bills. Hence, the difference between the small pricing errors of two highly correlated T-bill returns implies a large pricing error for the portfolio that has a long position in one T-bill and a short position in the other T-bill. Although the individual pricing errors of the short-maturity assets are comparable to those of the long-maturity assets, the higher correlation and lower variance of the short-maturity asset returns gives higher pricing errors for the short-maturity two-asset portfolios.

< Insert table 9 around here >

Table 9 is also useful to compare our results to those of Dai and Singleton (2001) and Duffee (2002). In section 4.3 we argued that our set of test restrictions is larger than the set used by Dai and Singleton (2001) and Duffee (2002), since we also focus on the term structure of unconditional expected bond returns. Table 9 shows that these additional restrictions are important: especially the pattern of unconditional expected returns on T-bills of different maturities leads to a rejection of the essentially affine models.

Overall, the conclusion is that under the assumption of a frictionless market up to three-factor completely and essentially affine yield term structure models do not seem to perform appropriately. One way to proceed is to turn to more-than-three-factor affine term structure models. However, in case of more-than-three-factor models, estimation becomes quickly much harder. A more interesting possibility is to examine models outside the class of (essentially) affine models, as is done by Ahn, Dittmar, and Gallant (2002), Duarte (2003), and Leippold and Wu (2003). Instead of this alternative, we choose to stick to the tractable class of affine models, and investigate whether the assumption of a frictionless market is too strong.

Therefore, we turn to the case *with transaction costs*.

< Insert table 10 around here >

In table 10 we present the results for the corresponding Wald test. We allow for transaction costs of τ (=$s/2$ in terms of equations (9) and (10)) basis points per holding period, assuming for simplicity that the transaction costs are the same for all transactions. We determine the critical transaction costs, defined as the amount of transaction costs for which the *p*-value of the Wald-test equals 0.05. For monthly holding periods, it follows that for relatively small amounts of transaction costs of less than 1 basis point, none of the preferred models is statistically rejected anymore. For quarterly holding periods larger transaction costs, up to 2.8 basis points for two-factor models and 2.6 basis points for the three-factor models, are required in order to avoid statistical rejection of the models. Because the monthly pricing errors are only very weakly correlated over time, the quarterly pricing errors are larger than the monthly

pricing errors and, therefore, also larger transaction costs are required to accept the model statistically. Confronting the critical transaction costs with the transaction costs as observed in the market, we see that the models only have difficulty in fitting the *Short Maturities Asset Set* with a quarterly holding period appropriately. Only for this case we find critical transaction costs (between 2 and 3 basis points) larger than the average of 1.5 basis points found in the data on T-bills. The evaluation of the term structure models thus becomes much more positive, than when judged on the basis of frictionless Euler restrictions: the difference between observed and model-implied (conditional) expected bond returns is not significant once we correct for market size transaction costs. As discussed earlier, Duffee (2002) notes that essentially affine models have problems in fitting both the time variation in expected returns and the time-varying behavior of volatilities. Our results show that allowing for transaction costs considerably softens the restrictions on expected (unconditional and conditional) bond returns. This may give more freedom in fitting the time-varying behavior of volatilities. Since we estimate the model parameters in a first step, this tradeoff is not directly observable in our parameter estimates. An analysis that combines the first-step ML estimation with the Euler restrictions in case of transaction costs, as used in the second step, is left for future research.

< Insert figures 1 and 2 around here >

In figures 1 and 2 we plot the SEBs as function of the transaction costs, distinguishing between the monthly and quarterly holding periods cases, respectively. Focusing first on the monthly holding period case, we see that figure 1 shows that the size of the SEB is below 0.1 at transaction costs of 1.6 basis point, which is economically rather small. In the frictionless case, extreme short and long positions in T-bills and bonds blow up the differences between pricing errors of T-bills and bonds so that standard test procedures reject the affine models. However, if small transaction costs are taken into account, these differences in pricing errors are not large enough to cause rejection of the models.

Next we turn to figure 2, showing the quarterly holding period case. Compared to the monthly holding period, larger transaction costs of more than 3 basis points are required in the quarterly holding period case to obtain a small SEB, in line with the findings for the Wald-tests. The figure shows that there is still

a strong influence of small transaction costs on the SEBs, although it is clearly less strong than for monthly holding periods.

Concluding, we see that the SEB-results are quite in line with the Wald-test results: with transaction costs of market size, the affine term structure models seem to perform reasonably well. Only for the *Short Maturities Asset Set* in case of a quarterly holding period, the restrictions implied by the models are rejected. As noted by Duffee (1996), the one- and two-month T-bill returns contain an idiosyncratic component, unrelated to returns on securities with longer maturities. His explanation for this idiosyncratic variation is market segmentation. This might (partially) explain the rejection of the models on the basis of the *Short Maturities Asset Set.*

4.6 Simulation study

The empirical results suggest that, when allowing for transaction costs, the affine yield models seem to be able to provide a reasonable description of the term structure of interest rates. In this subsection, we ask ourselves the following question: Assuming that the data are indeed generated by an affine yield factor model with transaction costs, can we explain why two and three-factor models are rejected in frictionless markets?

< Insert table 11 around here >

To answer this question, we perform a second simulation study. We simulate two-factor models, adding transaction costs of 1 basis point, and then estimate and test three-factor models assuming frictionless markets. The results are given in table 11. They indicate that a three-factor model cannot accurately fit returns generated by a two-factor model with transaction costs. A possible explanation for this is as follows. Assuming that the direction of trading is independent of the true state, transaction costs essentially add idiosyncratic factors to the bond returns, so that adding one extra common factor is unlikely to give correct pricing of bonds.

These simulation results show that the presence of transaction costs can lead to a rejection of three-factor affine models if one tests under the assumption of frictionless markets. In line with these

simulation results, we find empirically that three-factor models are rejected in case of frictionless markets, and that this apparent misspecification is resolved once we allow for modest transaction costs.

5 Summary and Conclusions

In this paper we analyze the bond pricing implications of both completely and essentially affine term structure models, with up to three factors, allowing for the presence of transaction costs. The goal of the paper is to assess the importance of incorporating market frictions for tests of bond pricing models.

Our tests focus on Euler equations, which can rewritten into restrictions on expected (conditional) returns on bonds of different maturities. By including the yield spread as conditioning variable, our tests include the implications of the affine models for the predictability of bond returns as studied by Dai and Singleton (2001) and Duffee (2002). However, our set of test restrictions is larger since we also include the term structure of unconditional expected bond returns.

We test two- and three-factor preferred models formally for different sizes of transaction costs, using a Wald-test, and we measure the size of misspecification of these models and analyze how sensitive the misspecification size is to the size of the transaction costs. Our analysis can be seen as an extension of Luttmer (1996), because we use the stronger specification error bound test, as opposed to the volatility bound that is used by Luttmer (1996). Also, Luttmer (1996) focuses on consumption-based asset pricing models, whereas we analyze models for the term structure of interest rates.

We find that, under the assumption of frictionless markets, completely and essentially affine yield models with up to three factors in general misprice T-bill and bond returns in a significant way. The term structure of average bond returns is less smooth than predicted by the model, so that long-short portfolios of near-maturity bonds are significantly mispriced. However, if we take transaction costs of market size into account, we find that the misspecification of the models disappears, in case of a monthly holding period. For quarterly holding periods and using market size transaction costs, the models fit long-maturity bond returns well, and are only rejected on the basis of short-maturity T-bill returns.

Appendix A: Identification Restrictions

Not all parameters in the (essentially) affine model in equations (2) and (3) can be identified. To ensure econometric identification of the model, we impose the same normalizations as in Dai and Singleton (2000) and Duffee (2002). For the sake of completeness, we describe the identification restrictions in this appendix.

We start by discussing the case of completely affine models (2). Dai and Singleton (2000) show that the class of N-factor completely affine models consists of $N+1$ non-nested subclasses of models. The classification is based on the number of factors m that enter the conditional variance of the factors. The associated subfamilies are denoted $A_0(N),..,A_m(N),..,A_N(N)$.

We define $K = I_N - \Lambda$ in order to maintain a similar notation as in Dai and Singleton (2000). We partition a_t as $a_t = (a_t^{B/}, a_t^{D/})$, where a_t^B is m by 1 and a_t^D is $(N$-$m)$ by 1. We now list the restrictions on the parameters that lead to identification of the remaining parameters for the $A_m(N)$ subclass. First of all, if m=0, K is upper or lower triangular, and for m>0

$$K = \begin{bmatrix} K_{mxm}^{BB} & 0_{mx(N-m)} \\ K_{(N-m)xm}^{DB} & K_{(N-m)x(N-m)}^{DD} \end{bmatrix}. \tag{A.1}$$

In addition, for all m=$0,...,N,$ we normalize

$$\mu = \begin{bmatrix} \mu_{mx1}^{B} \\ 0_{(N-m)x1} \end{bmatrix},$$

$$\Sigma = I,$$

$$\alpha = \begin{bmatrix} 0_{mx1} \\ 1_{(N-m)x1} \end{bmatrix},$$

$$B = (\beta_1,...,\beta_n) = \begin{bmatrix} I_{mxm} & B_{mx(N-m)}^{BD} \\ 0_{(N-m)xm} & 0_{(N-m)x(N-m)} \end{bmatrix}, \tag{A.2}$$

and impose the following parameter restrictions

$$\varphi_i \geq 0, \quad m+1 \leq i \leq N,$$

$$\sum_{j=1}^{m} \mathrm{K}_{ij}\mu_j > 0, \quad 1 \leq i \leq m,$$

$$\mathrm{K}_{ij} \leq 0, \quad 1 \leq j \leq m, \ j \neq i,$$

$$\mu_i \geq 0, \quad 1 \leq i \leq m,$$

$$\mathrm{B}_{ij} \geq 0, \ 1 \leq i \leq m, \ m+1 \leq j \leq N.$$

(A.3)

Next, we discuss the restrictions in essentially affine models, thus allowing for a nonzero Γ matrix in (2) and (3). In addition to the restrictions listed above, Duffee (2002) shows that in an essentially affine $EA_m(N)$ model the following normalization renders identification of the remaining parameters

$$\Gamma = \begin{bmatrix} 0_{mxN} \\ \Gamma^D_{(N-m)xN} \end{bmatrix}.$$

(A.4)

This implies that the class of $EA_N(N)$ models coincides with the $A_N(N)$ class.

Appendix B: Kalman Filter ML Estimation

In this appendix, we briefly describe the Kalman filter estimation of affine term structure models. For a more detailed exposition, we refer to Duan and Simonato (1999) and De Jong (2000).

The Kalman filter state space model is characterized by the transition equation and the measurement equation. The affine yield models in equation (2) and (3) provide the following transition equation for the factors

$$a_{t+1} = \mu + \Lambda(a_t - \mu) + \Sigma \xi_{t+1}.$$

(B.1)

Given the normality of ξ_{t+1}, the conditional distribution of the factors a_{t+1} is multivariate normal, with conditional expectation and covariance matrix given by

$$E[a_{t+1}|a_t] = \mu + \Lambda(a_t - \mu), \quad V[a_{t+1}|a_t] = \Sigma\, V(a_t)\Sigma'. \tag{B.2}$$

These expressions are simpler than those presented in De Jong (2000), because we use a discrete-time affine model.

The second part of a state-space model is the measurement equation. We use zero-coupon interest rates to construct this equation. Let $r_t = (r_{n(1)t},...,r_{n(K)t})'$ denote a K-dimensional vector with the time-t interest rates of different maturities $n(1),...,n(K)$, with $K>N$. From equation (5) we then obtain, introducing a K-dimensional vector of measurement errors ε_t,

$$r_t = \frac{1}{n}A + \frac{1}{n}B'a_t + \varepsilon_t. \tag{B.3}$$

Here $A = (A_{n(1)},...,A_{n(K)})'$ and $B = (B_{n(1)},...,B_{n(K)})$. The vector ε_t is assumed to be i.i.d. and normally distributed with

$$E[\varepsilon_t] = 0, \quad V[\varepsilon_t] = diag(\sigma_1^2,....,\sigma_K^2). \tag{B.4}$$

The errors are thus assumed to be independent across maturities. This completes the state-space model.

Given the conditional normality of ε_t and ξ_{t+1}, Maximum Likelihood (ML) yields consistent and efficient parameter estimates (Hamilton (1994)). We refer to De Jong (2000) for the Kalman filter equations that are needed to construct the Likelihood function. We assume that all factors follow stationary processes, so that we can use the unconditional expectation and (co)variances of the factors to initiate the Kalman filter.

A well-known issue with Kalman Filter ML estimation of affine term structure models is that the conditional variance of ξ_{t+1} depends on the unknown values a_t, which makes the ML estimator based on the Kalman Filter strictly speaking inconsistent. Simulation evidence by Duan and Simonato (1999) and De Jong (2000) shows that the resulting biases are very small for reasonable sample sizes. One can obtain consistent estimates by using the Efficient Method of Moments (EMM, Gallant and Tauchen (1996)), combined with the semi-nonparametric (SNP) method of Gallant and Tauchen (1992). Duffee and

Stanton (2000) provide a comparison of EMM/SNP estimation and Kalman Filter ML estimation. They report important small-sample biases for the EMM/SNP method, and conclude that 'for reasonable sample sizes, the results strongly support the choice of the Kalman filter'.

Appendix C: Limit Distribution of Wald Test

We first discuss how our set-up fits into the framework of Kodde and Palm (1986) (KP from now on) in case the SDF is fully observable. After this, we discuss the required modifications when the SDF includes unknown parameters that are estimated in a preliminary round.

Our case fits in KP, case 2, with $h(\theta) = h_2(\theta) = 0, q = 0$. In our case, the definition of θ is given by

$$
\begin{aligned}
\theta &= (\theta_1', \theta_2')' \\
\theta_1 &= (\theta_{11}, ..., \theta_{1n})' \\
\theta_2 &= (\theta_{21}, ..., \theta_{2n})' \\
\theta_{1i} &= E[Y_{t+1} x_{i,t+1}], \ i = 1, ..., n \\
\theta_{2i} &= E[q_{it}], \ i = 1, ..., n
\end{aligned}
\tag{C.1}
$$

The function $h(\theta) = h_2(\theta)$ is defined by

$$
\begin{aligned}
h(\theta) &= (h_1(\theta)', ..., h_n(\theta)')' \\
h_i(\theta) &= (h_i^a(\theta) \ h_i^b(\theta))', \ i = 1, ..., n \\
h_i^a(\theta) &= \theta_{1i}/\theta_{2i} - 1/\tau^s, \ i = 1, ..., n \\
h_i^b(\theta) &= -\theta_{1i}/\theta_{2i} + 1/\tau^l, \ i = 1, ..., n
\end{aligned}
\tag{C.2}
$$

Then we can formulate the null hypothesis in the main text in terms of the notation of KP as H_0: $h(\theta) \geq 0$.

However, in this set-up the components $h_i^a(\theta)$ and $h_i^b(\theta)$, $i = 1, ..., n$ are not independent, as required by KP. We take a local point of view, which means that for each $i = 1, ..., n$ at most one of the two possibilities $h_i^a(\theta) \geq 0$ and $h_i^b(\theta) \geq 0$ will be relevant. Thus, when deriving the limit distribution of the test statistic under the

null, we only have to include the relevant $h_i^j(\theta), j \in \{a,b\}, i = 1,...,n$, which we stack into an n-dimensional vector function $\tilde{h}(\theta)$.

As estimator for θ we take

$$\hat{\theta} = (\hat{\theta}_1', \hat{\theta}_2')'$$

$$\hat{\theta}_1 = (\hat{\theta}_{11},...,\hat{\theta}_{1n})'$$

$$\hat{\theta}_2 = (\hat{\theta}_{21},...,\hat{\theta}_{2n})'$$

$$\hat{\theta}_{1i} = \frac{1}{T}\sum_t Y_{t+1} x_{i,t+1}, \ i = 1,...,n$$

$$\hat{\theta}_{2i} = \frac{1}{T}\sum_t q_{it}, \ i = 1,...,n$$

(C.3)

Obviously, using an appropriate version of the Central Limit Theorem, we have

$$\sqrt{T}(\hat{\theta} - \theta) \to^d N(0, V(\theta))$$

(C.4)

where $V(\theta)$ can be partitioned in accordance with $\theta = (\theta_1' \ \theta_2')'$ as

$$V(\theta) = \begin{pmatrix} V_{11}(\theta) & V_{12}(\theta) \\ V_{21}(\theta) & V_{22}(\theta) \end{pmatrix}$$

(C.5)

Define $H_{1i}(\theta) = 1/\theta_{2i}$, $H_{2i}(\theta) = -\theta_{1i}/\theta_{2i}^2$. Then

$$\frac{\partial h_i^a(\theta)}{\partial \theta_{1i}} = H_{1i}(\theta), \ \frac{\partial h_i^b(\theta)}{\partial \theta_{1i}} = -H_{1i}(\theta)$$

$$\frac{\partial h_i^a(\theta)}{\partial \theta_{2i}} = H_{2i}(\theta), \ \frac{\partial h_i^b(\theta)}{\partial \theta_{2i}} = -H_{2i}(\theta)$$

(C.6)

So, the difference between the ^{a}- and the ^{b}-version is simply the sign. Let $H_1(\theta) = diag(H_{11}(\theta),...,H_{1n}(\theta))$ and $H_2(\theta) = diag(H_{21}(\theta),...,H_{2n}(\theta))$, and let $\tilde{H}_1(\theta)$ and $\tilde{H}_2(\theta)$ denote the versions of $H_1(\theta)$ and $H_2(\theta)$ with the relevant ^{a}- or the ^{b}-version included (following our local point of view). Applying the delta-method, we find

$$\sqrt{T}(\tilde{h}(\hat{\theta}) - \tilde{h}(\theta)) = (\tilde{H}_1(\theta), \tilde{H}_2(\theta))\sqrt{T}(\hat{\theta} - \theta) + o_p(1) \tag{C.7}$$

so that the limit distribution of $\sqrt{T}(\tilde{h}(\hat{\theta}) - \tilde{h}(\theta))$ is the normal distribution with mean 0 and covariance matrix

$$(\tilde{H}_1(\theta), \tilde{H}_2(\theta)) \begin{pmatrix} V_{11}(\theta) & V_{12}(\theta) \\ V_{21}(\theta) & V_{22}(\theta) \end{pmatrix} \begin{pmatrix} \tilde{H}_1(\theta) \\ \tilde{H}_2(\theta) \end{pmatrix} \tag{C.8}$$

Due to the structure of $\tilde{H}_1(\theta)$ and $\tilde{H}_2(\theta)$ it is easy to see that this is equal to

$$(H_1(\theta), H_2(\theta)) \begin{pmatrix} V_{11}(\theta) & V_{12}(\theta) \\ V_{21}(\theta) & V_{22}(\theta) \end{pmatrix} \begin{pmatrix} H_1(\theta) \\ H_2(\theta) \end{pmatrix} \tag{C.9}$$

which is the Σ-matrix used by KP. This completes the discussion of how our set-up fits in KP, in case the SDF does not depend upon unknown parameters.

When the SDF depends upon unknown parameters that are estimated in a preliminary round we have the following modification of the set-up presented in KP. First, we have

$$\theta_1 = \theta_1(\alpha) = (\theta_{11}(\alpha),...,\theta_{1n}(\alpha))^{/}$$
$$\theta_{1i}(\alpha) = E[Y_{t+1}(\alpha)x_{i,t+1}], \quad i=1,...,n \tag{C.10}$$

Since α is unknown, we estimate it in a preliminary round by $\hat{\alpha}$, satisfying

$$\sqrt{T}(\hat{\alpha} - \alpha) = \frac{1}{\sqrt{T}}\sum_t \psi_t(\alpha) + o_p(1) \tag{C.11}$$

with $E(\psi_t(\alpha)) = 0$, $E(\psi_t(\alpha)^2) < \infty$. We estimate $\theta_{1i}(\alpha)$, $i=1,...,n$, by

$$\tilde{\theta}_{1i} = \frac{1}{T}\sum_t Y_{t+1}(\hat{\alpha})x_{i,t+1}, \quad i=1,...,n \tag{C.12}$$

We now illustrate how the limit distribution of $\tilde{\theta}_{1i}$ can be derived. The general case for $\tilde{\theta} = (\tilde{\theta}_1^{/}, \hat{\theta}_2)^{/}$,

with $\tilde{\theta}_1 = (\tilde{\theta}_{11},...,\tilde{\theta}_{1n})'$, then follows straightforwardly. We have

$$\sqrt{T}\Big(\frac{1}{T}\sum_t Y_{t+1}(\hat{\alpha})x_{i,t+1} - E(Y_{t+1}(\alpha)x_{i,t+1})\Big)$$

$$= \sqrt{T}\Big(\frac{1}{T}\sum_t Y_{t+1}(\alpha)x_{i,t+1} - E(Y_{t+1}(\alpha)x_{i,t+1})\Big) + \frac{1}{T}\sum_t \frac{\partial Y_{t+1}(\alpha)x_{i,t+1}}{\partial\alpha}\sqrt{T}(\hat{\alpha}-\alpha) + o_p(1)$$

$$= \sqrt{T}\Big(\frac{1}{T}\sum_t Y_{t+1}(\alpha)x_{i,t+1} - E(Y_{t+1}(\alpha)x_{i,t+1})\Big) + \frac{1}{T}\sum_t \frac{\partial Y_{t+1}(\alpha)x_{i,t+1}}{\partial\alpha'}\frac{1}{\sqrt{T}}\sum_t \psi_t(\alpha) + o_p(1)$$

$$\hspace{9cm}\text{(C.13)}$$

$$= \Big(1,\ \frac{\partial Y_{t+1}(\alpha)x_{i,t+1}}{\partial\alpha'}\Big)\sqrt{T}\frac{1}{T}\sum_t \begin{pmatrix} Y_{t+1}(\alpha)x_{i,t+1} - E(Y_{t+1}(\alpha)x_{i,t+1}) \\ \psi_t(\alpha) \end{pmatrix}$$

from which the limit distribution of $\frac{1}{T}\sum_t Y_{t+1}(\hat{\alpha})x_{i,t+1}$ follows without further complications. Given the limit distribution of $\tilde{\theta}$ the limit distribution of $h(\tilde{\theta})$ follows similarly to the case without initial parameter estimation.

References

Ahn, C., Dittmar, R.F., Gallant, A.R., 2002, Quadratic Term Structure Models: Theory and Evidence, *Review of Financial Studies*, 15, 1, 243-288.

Ahn, C., Dittmar, R.F., Gallant, A.R., Gao, B., 2002, Purebred or Hybrid? Reproducing the Volatility in Term Structure Dynamics, *Journal of Econometrics*, forthcoming.

Babbs, S.H., Nowman, K.B., 1999, Kalman Filtering of Generalized Vasicek Term Structure Models, *Journal of Financial and Quantitative Analysis,* 34, 115-130.

Backus, D., Foresi, S., Mozumdar, A., Wu, L., 2001, Predictable changes in yields and forward rates, *Journal of Financial Economics*, 59, 281-311.

Backus, D., Foresi, S., Telmer, C.I., 2001, Affine term structure models and the forward premium anomaly, *The Journal of Finance*, 56, 279-304.

Backus, D.K., Zin, S.E., 1994, Reverse Engineering the Yield Curve, Working Paper NBER.

Bansal, R., 1997, An Exploration of the Forward Premium Puzzle in Currency Markets, *Review of Financial Studies,* 10, 131-166.

Bansal, R., Zhou, H., 2002, Term Structure of Interest Rates with Regime Shifts, *Journal of Finance*, 57, 1997-2043.

Brown, R.H., Schaefer, S.M., 1994, The Term Structure of Real Interest Rates and the CIR Model,

Journal of Financial Economics, 35, 3-42.

Boudoukh, J., Richardson, M., Stanton, R., Whitelaw, R.F., 1999, The Stochastic Behavior of Interest Rates: Implications from a Multifactor, Nonlinear Continuous-time Model, Working Paper, New York University.

Campbell, J.Y., Lo, A., MacKinley, A., 1997, *The Econometrics of Financial Markets,* Princeton University Press.

Campbell, J.Y., Shiller, R.J., 1991, Yield Spreads and Interest Rate Movements: A Bird's Eye View, *Review of Economic Studies*, 58, 495-514.

Chakravarty, S., Sarkar, A., 1999, Liquidity in U.S. Fixed Income Markets: A Comparison of the Bid-Ask Spread in Corporate, Government and Municipal Bond Markets, Working Paper, Purdue University.

Chen, R., Scott, L., 1993, ML Estimation for a Multi-Factor Equilibrium Model of the Term Structure, *Journal of Fixed Income,* December.

Cochrane, J.H., 1996, A Cross-sectional Test of an Investment-based Asset Pricing Model, *Journal of Political Economy*, 104, 572-621.

Cochrane, J.H., 2001, *Asset Pricing*, Princeton University Press.

Cox, J., Ingersoll, J., Ross, S., 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-408.

Dai, Q., Singleton, K.J., 2000, Specification Analysis of Affine Term Structure Models, *Journal of*

Finance 55, 1943-1978.

Dai, Q., Singleton, K.J., 2002, Expectations Puzzles, Time-Varying Risk Premia, and Dynamic Models of the Term Structure, *Journal of Financial Economics,* 63, 415-441.

Dai, Q., Singleton, K.J., 2003, Term Structure Dynamics in Theory and Reality, *Review of Financial Studies,* 16, 631-678.

De Jong, F., 2000, Time-series and Cross-section Information in Affine Term Structure Models, *Journal of Economics and Business Statistics*, 18, 300-318.

Duan, J.C., Simonato, J.G., 1999, Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter, *Review of Quantitative Finance and Accounting*, 13, 111-135.

Duarte, J., 2003, Evaluating an Alternative Risk Preference in Affine Term Structure Models, *Review of Financial Studies*, forthcoming.

Duffee, G., 1996, Idiosyncratic Variation of Treasury Bill Yield Spreads, *Journal of Finance*, 51, 527-552.

Duffee, G.R., 2002, Term premia and interest rate forecasts in affine models, *The Journal of Finance*, 57, 405-443.

Duffee G.R., Stanton, R.H., 2000, Estimation of Dynamic Term Structure Models, Working paper, Haas School of Business, U.C. Berkeley.

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

Fama, E., Bliss, R., 1987, The Information in Long-Maturity Forward Rates, *American Economic Review*, 77(4), 680-692.

Ferson, W.E., Siegel, A.F., 2003, Stochastic Discount Factor Bounds with Conditioning Information, *The Review of Financial Studies*, 16, 567-595.

Gallant A.R., Tauchen, G.E., 1992, A Nonparametric Approach to Nonlinear Time Series Analysi: Estimation and Simulation, in: D. Brillinger, P. Caines, E. Parzen, M> Rosenblatt, and M.S. Taqqu (eds.), *New Directions in Times Series Analysis, Part II*, 71-92, Springer-Verlag, New York.

Gallant A.R., Tauchen, G.E., 1996, Which Moments to Match?, *Econometric Theory*, 12, 657-681.

Gibbons, M.R., Ramaswamy, K., 1993, A Test of the CIR Model of the Term Structure, *Review of Financial Studies,* 6, 619-658.

Hansen, L.P., 1982, Large Sample Properties of Generalized Methods of Moments Estimators, *Econometrica,* 50, 1029-1054.

Hansen, L.P., Jagannathan, R., 1991, Implications of Security Market Data for Models of Dynamic Economies, *Journal of Political Economy,* 99, 225-262.

Hansen, L.P., Jagannathan, R., 1997, Assessing Specification Errors in Stochastic Discount Factor Models, *Journal of Finance,* 52, 557-590.

Hansen, L.P., Heaton, J.C., Luttmer, E., 1995, Econometric Evaluation of Asset Pricing Models, *Review of Financial Studies,* 8, 237-274.

He, H., Modest, D.M., 1995, Market Frictions and Consumption-Based Asset Pricing, *Journal of Political Economy,* 101, 94-117.

Jouini, E., Kalal, H., 1995, Martingales, Arbitrage and Securities Markets with Transaction Costs, *Journal of Economic Theory,* 66, 178-197.

Kan, R., Zhou, G., 1999, A Critique of the Stochastic Discount Factor Methodology, *Journal of Finance* 54, 1221-1248.

Kodde, D.A., Palm, F.C., 1986, Wald Criteria for Jointly Testing Equality and Inequality Restrictions, *Econometrica,* 54, 1243-1248.

Leippold, M., Wu, L., 2003, Design and Estimation of Quadratic Term Structure Models, *European Finance Review*, 7(1), 47-73.

Luttmer, E., 1996, Asset Pricing in Economies with Frictions, *Econometrica,* 64, 1439-1467.

Newey, W.K., West, K.D., 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica,* 55, 703-708.

Pearson, N.D., Sun, T.S., 1994, Exploiting the Conditional Density in Estimating the Term Structure: an Application of the CIR model, *Journal of Finance,* 49, 1279-1304.

Stambaugh, R., 1988, The Information in Forward Rates, *Journal of Financial Economics,* 21, 41-70.

Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics,* 5, 177-188.

Wolak, F.A., 1989, Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models, *Econometric Theory,* 5, 1-35.

Wolak, F.A., 1991, The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models, *Econometrica,* 59, 981-995.

Tables

Table 1. Properties of Interest Rates.

The sample moments are calculated for interest rates from the CRSP Fama Files, which contains monthly data for the period 1972-1997. Interest rates are expressed on a yearly basis.

Maturity in Months	Mean	St. Deviation	Autocorrelation
1	6.61%	2.67%	0.96
3	7.03%	2.78%	0.97
12	7.50%	2.64%	0.97
24	7.78%	2.49%	0.98
36	7.96%	2.38%	0.98
48	8.11%	2.30%	0.98
60	8.20%	2.24%	0.98

Table 2. Properties of Holding Returns.

The table contains sample moments of monthly holding returns on T-bills with maturities of 1, 3 ,6 and 9 months and monthly holding returns on portfolios of bonds with maturities in a certain maturity interval, which are calculated using CRSP data for 1972-1997.

Maturity in Months	Mean	St. Deviation	Autocorrelation
1	0.56%	0.23%	0.95
3	0.61%	0.27%	0.81
6	0.63%	0.36%	0.50
9	0.65%	0.49%	0.36
24-36	0.69%	1.20%	0.17
48-60	0.70%	1.67%	0.15
60-120	0.73%	2.07%	0.14
> 120	0.78%	2.92%	0.12

Table 3. Bid-Ask Spreads of T-bill Prices.

Bid-ask spreads are in basis points and calculated by dividing the difference between bid and ask prices by the mid price. Data on bid and ask prices come from the CRSP T-bill dataset.

Maturity	Average Bid-Ask Spread, 1972-1997.	Standard Deviation, 1972-1997.	Average Bid-Ask Spread, 1987-1997.	Standard Deviation, 1987-1997.
1 month	1.3 bp	1.4 bp	0.4 bp	0.3 bp
3 months	2.1 bp	2.1 bp	0.5 bp	0.2 bp
6 months	4.1 bp	4.0 bp	1.1 bp	0.2 bp
9 months	5.9 bp	5.3 bp	1.6 bp	0.3 bp

Table 4. Kalman Filter ML Estimation Results: Preferred Models.

The table contains the results of Kalman Filter ML estimation of discrete-time, monthly, two-factor and three-factor affine models. In appendix B, details on Kalman Filter ML estimation are provided. In brackets the standard errors are given. Also presented are the mean and variance of the implied stochastic discount factor, as well as the mean absolute yield error for the fitted interest rates. All parameters are expressed on a monthly basis.

Parameter	$A_0(2)$ model	$A_1(3)$ model	$EA_1(3)$ model
φ_0	0.006 (0.003)	0.003 (0.002)	0.002 (0.002)
$100\ \varphi_{1,1}$	0.036 (0.014)	0.002 (0.002)	0.002 (0.003)
$100\ \varphi_{1,2}$	0.047 (0.028)	0.029 (0.021)	0.027 (0.022)
$100\ \varphi_{1,3}$	-	0.058 (0.049)	0.060 (0.045)
μ_1	-	1.238 (0.982)	1.412 (1.083)
Λ_{11}	0.998 (0.023)	0.999 (0.019)	0.999 (0.019)
Λ_{22}	0.925 (0.048)	0.673 (0.174)	0.713 (0.187)
Λ_{21}	0.006 (0.004)	-	-
Λ_{33}	-	0.949 (0.099)	0.933 (0.134)
Λ_{31}	-	0.009 (0.007)	0.011 (0.008)
$\beta_{2,1}$	-	0.011 (0.006)	0.017 (0.013)
$100\ \gamma_1$	-1.082 (1.277)	-0.125 (0.345)	-
$100\ \gamma_2$	-2.636 (0.914)	-0.541 (0.246)	-0.384 (0.232)
$100\ \gamma_3$	-	-0.520 (0.427)	-0.355 (0.319)
$100\ \Gamma_{21}$	-	-	-0.059 (0.066)
$100\ \Gamma_{32}$	-	-	0.094 (0.086)
$100\ \Gamma_{33}$	-	-	0.035 (0.012)
Mean of SDF	0.9958	0.9942	0.9952
St.Dev. of SDF	0.0313	0.0540	0.0813
Mean Absolute yield error	8.75 basis points	6.62 basis points	6.02 basis points

Table 5. Pricing Error Correlations across Preferred Models.

For each asset in the all maturities asset-set, the correlation between the monthly pricing errors of two models is calculated. The table presents the average of these correlations over all assets, for each pair of the preferred models. The table also contains for each model the average pricing error, averaged over all assets, as well as the t-ratio of this average.

	Avg. Pricing Error [t-ratio]	$A_0(2)$	$A_1(3)$	$EA_1(3)$
$A_0(2)$	-0.0035 [1.75]	1		
$A_1(3)$	-0.0039 [1.88]	0.672	1	
$EA_1(3)$	-0.0020 [0.92]	0.538	0.772	1

Table 6. Simulation of Three-Factor Model with Transaction Costs:
Test of Three-Factor and Two-Factor Models

For each benchmark three-factor model, a monthly time series (25 years length) of interest rates and bond prices is simulated, adding transaction costs of 1.0 basis point to bond prices. In total, 1000 simulations are performed. Next, three-factor (Panel A) and two-factor (Panel B) affine models are tested on these simulated data, using the Wald test and the SEB, both for the assumption of frictionless markets and allowing for transaction costs. We use a 5% critical value for the Wald test. Finally, the table reports the amount of transactions costs that should be incorporated in the Wald test to obtain a rejection rate of 95% for the two-factor model.

Panel A: Tests of Three-Factor Models			
Benchmark Model	$A_0(3)$	$A_1(3)$	$A_3(3)$
Tested Model	$A_0(3)$	$A_1(3)$	$A_3(3)$
Tests in case of frictionless markets			
Percentage rejections Wald test	11.5%	12.2%	13.0%
Median SEB	0.127	0.142	0.133
Tests in case of transaction costs: 1.0 bp			
Percentage rejections Wald test at transaction costs of 1.0 bp	4.0%	3.8%	3.7%
Median SEB at transaction costs of 1.0 bp	0.035	0.042	0.040
Panel B: Tests of Two-Factor Models			
Benchmark Model	$A_0(3)$	$A_1(3)$	$A_3(3)$
Tested Model	$A_0(2)$	$A_1(2)$	$A_2(2)$
Tests in case of frictionless markets			
Percentage rejections Wald test	97%	98%	97%
Median SEB	0.487	0.511	0.505
Tests in case of transaction costs: 1.0 bp			
Percentage rejections Wald test at transaction costs of 1.0 bp	65.6%	56.8%	61.5%
Median SEB at transaction costs of 1.0 bp	0.227	0.232	0.245
Transaction costs at which rejection rate equals 95%	0.1 bp	0.1 bp	0.1 bp

Table 7.Wald-test for Preferred Two-Factor and Three-Factor Models in Frictionless Markets.

The table reports Wald test results for two- and three-factor affine models and monthly and quarterly holding periods. To calculate the asymptotic covariance matrices for the quarterly holding period, we use the Newey-West method with 2 lags to correct for the overlapping nature of the returns.

	$A_0(2)$	$A_1(3)$	$EA_1(3)$
Monthly holding period			
p-value: *Short-maturities*	0.000	0.000	0.000
p-value: *Long-maturities*	0.033	0.089	0.047
p-value: *All-maturities*	0.000	0.000	0.000
Quarterly holding period			
p-value: *Short-maturities*	0.000	0.000	0.000
p-value: *Long-maturities*	0.028	0.029	0.012
p-value: *All-maturities*	0.000	0.000	0.000

Table 8. SEB for Preferred Two-Factor and Three-Factor Models in Frictionless Markets.

The table reports the SEB for two-factor and three-factor affine models. Asymptotic standard errors of the SEB are given in brackets. To calculate the asymptotic covariance matrices for the quarterly holding period, we use the Newey-West method with 2 lags to correct for the overlapping nature of the returns.

	$A_0(2)$	$A_1(3)$	$EA_1(3)$
Monthly holding period			
SEB: *Short-maturities*	0.613 (0.057)	0.622 (0.067)	0.598 (0.072)
SEB: *Long-maturities*	0.195 (0.053)	0.189 (0.053)	0.191 (0.057)
SEB: *All-maturities*	0.642 (0.063)	0.638 (0.073)	0.617 (0.081)
Quarterly holding period			
SEB: *Short-maturities*	0.948 (0.088)	0.947 (0.082)	0.921 (0.111)
SEB: *Long-maturities*	0.342 (0.068)	0.327 (0.057)	0.352 (0.070)
SEB: *All-maturities*	1.165 (0.092)	1.157 (0.100)	1.150 (0.138)

Table 9. Normalized Absolute Pricing Errors of Long-Short Portfolios: Essentially Affine Three-Factor Model.

The table contains monthly absolute pricing errors of the essentially affine $EA_1(3)$ model, for one- and two-asset portfolios. For each T-bill, the long-short portfolio refers to a portfolio of the particular T-bill and the 1-month T-bill. For each bond, the long-short portfolio refers to a portfolio in the particular bond and the 2-3 year maturity bond. For these long-short portfolios, the multiplier-vector or weight-vector λ in (17) always contains a positive and an equally large negative element. The portfolio weights are normalized as in equation (17). In brackets, standard errors of the pricing errors are presented.

	Normalized Pricing Error Individual Assets	Normalized Pricing Error Long-Short Portfolios	Normalized Pricing Error Individual Assets	Normalized Pricing Error Long-Short Portfolios
	Unconditional Returns		Conditional Returns	
T-bill 1-month	0.0030 (0.0020)	-	0.0041 (0.0018)	-
T-bill 3-months	0.0025 (0.0020)	0.289 (0.045)	0.0037 (0.0018)	0.297 (0.067)
T-bill 6-months	0.0023 (0.0019)	0.092 (0.042)	0.0032 (0.0021)	0.121 (0.061)
T-bill 9-months	0.0021 (0.0019)	0.044 (0.034)	0.0026 (0.0022)	0.067 (0.046)
Bond 2-3 years	0.0018 (0.0016)	-	0.0027 (0.0019)	-
Bond 4-5 years	0.0018 (0.0015)	0.066 (0.054)	0.0025 (0.0019)	0.045 (0.048)
Bond 5-10 years	0.0015 (0.0016)	0.038 (0.037)	0.0021 (0.0020)	0.019 (0.043)
Bond >10 years	0.0012 (0.0017)	0.025 (0.035)	0.0022 (0.0021)	0.010 (0.039)

Table 10. Critical Transaction Costs for Two-Factor and Three-Factor Models.

The critical transaction costs are defined as the amount of transaction costs for which the Wald p-value is equal to 5%. Transaction costs are relative to the price and presented in basis points. The table presents results for monthly and quarterly holding periods, and two- and three-factor affine models.

	$A_0(2)$	$A_1(3)$	$EA_1(3)$
Monthly holding period			
Short-maturities	0.8 bp	0.7 bp	0.8 bp
Long-maturities	0.4 bp	0.4 bp	0.3 bp
All-maturities	0.9 bp	0.8 bp	0.8 bp
Quarterly holding period			
Short-maturities	2.3 bp	2.2 bp	2.4 bp
Long-maturities	0.9 bp	0.9 bp	0.8 bp
All-maturities	2.8 bp	2.6 bp	2.6 bp

Table 11. Simulation of Two-Factor Model with Transaction Costs: Tests of Three-Factor Models in Frictionless Markets

For each benchmark two-factor model, a monthly time series (25 years length) of interest rates and bond prices is simulated, adding transaction costs of 1 basis point to bond prices. Next, three-factor affine models are tested on these simulated data, using the Wald test and the SEB, and assuming frictionless markets. In total, 1000 simulations are performed. We use a 5% critical value for the Wald test.

Benchmark Model	$A_0(2)$	$A_1(2)$	$A_2(2)$
Tested Model	$A_0(3)$	$A_1(3)$	$A_3(3)$
Percentage rejections Wald test	95.6%	98.2%	98.8%
Median SEB	0.407	0.429	0.461

Figure 1. Monthly SEB for Two-Factor and Three-Factor Models. The graph shows the SEB for different sizes of transaction costs, for the preferred two-factor and three-factor affine models, in case of monthly holding periods and the all-maturities asset set.

Figure 2. Quarterly SEB for Two-Factor and Three-Factor Models. The graph shows the SEB for different sizes of transaction costs, for the preferred two-factor and three-factor affine models, in case of quarterly holding periods and the all-maturities asset set.
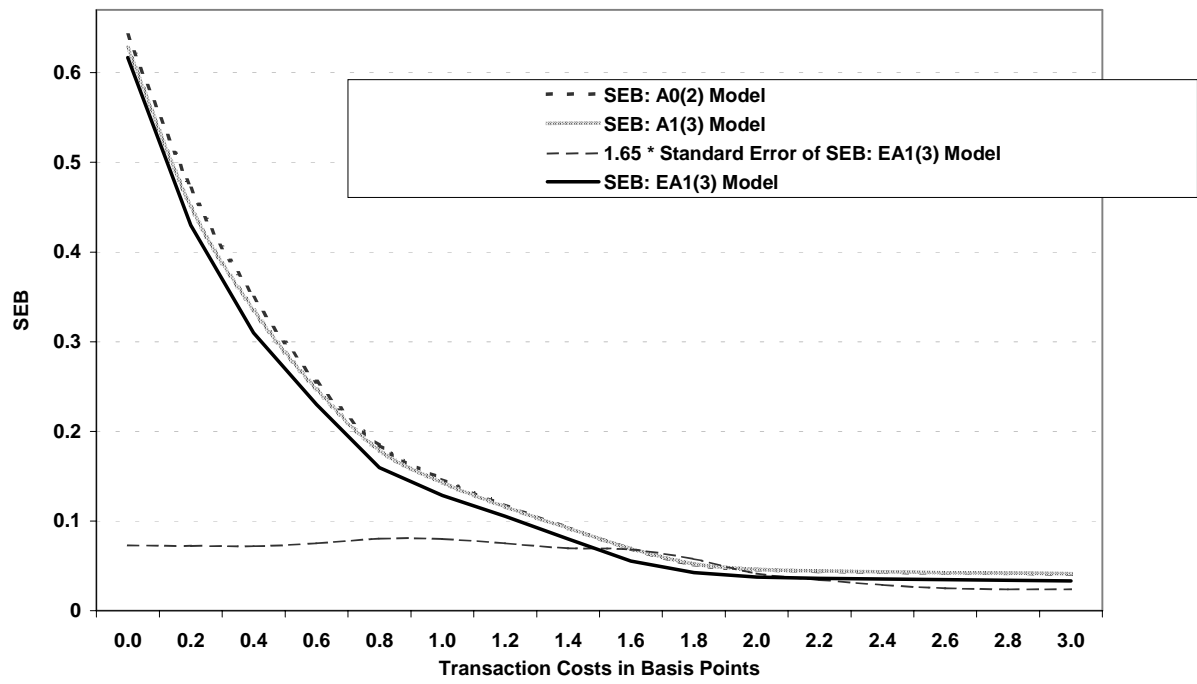
Figure 1

Figure 2