Risk Premia, Volatilities, and Sharpe Ratios in a Nonlinear Term Structure Model*

Peter Feldhütter†
London Business School

Christian Heyerdahl-Larsen‡
London Business School

Philipp Illeditsch§
The Wharton School

December 17, 2013

Abstract

We introduce a new reduced form term structure model where the short rate and market prices of risk are nonlinear functions of Gaussian state variables but yields are nevertheless given in closed form. Empirically, our three-factor nonlinear Gaussian model matches both the time-variation in expected excess returns and yield volatilities of U.S. Treasury bonds. During low volatility periods Treasury bonds are more attractive investments than standard Gaussian models predict. A significant part of expected excess returns in the nonlinear model is not explained by a linear combination of yields. This suggests that more information about expected excess returns is contained in the yield curve than previously thought, but in a nonlinear way.

Keywords: Nonlinear term structure models, Gaussian term structure models, time-varying expected excess returns, stochastic volatility, Sharpe ratios, hidden factors.

JEL Classification: D51, E43, E52, G12.

*We would like to thank Kerry Back; Greg Bauer; David Chapman; Mikhail Chernov; Alex David; Greg Duffee; Paul Ehling; Rodrigo Guimaraes; Burton Hollifield; Scott Joslin; Ralph Koijen; Andreas Pick; Christian Opp; Giuliano De Rossi; Ivan Shaliastovich; Amir Yaron; seminar participants at the London Business School, the Wharton School, the University of Melbourne; and participants at the Tripartite Conference held at Wharton, the World Finance Conference in Cyprus, the UBC Summer conference in Vancouver, and the Inquire UK conference for helpful comments and suggestions.

†London Business School, Regent's Park, London, NW1 4SA, pfeldhutter@london.edu

‡London Business School, Regent's Park, London, NW1 4SA, cheyerdahllarsen@london.edu

§The Wharton School, University of Pennsylvania, pille@wharton.upenn.edu

I Introduction

Investments in U.S. Treasury bonds require an assessment of both risk and reward. Existing term structure models are not flexible enough to capture the time series variation of both conditional first and second moments of yields.[1] Time variation in bond volatility is often sacrificed and affine Gaussian models are used because of their analytical tractability.[2] We propose a new reduced form term structure model where the short rate, market prices of risk, and bond volatilities are nonlinear functions of Gaussian state variables. We provide closed-form solutions for bond prices and since the factors are Gaussian our nonlinear model is as tractable as the standard Gaussian model.

We use U.S. Treasury bond data from 1952 to 2011 to estimate a standard Gaussian and a nonlinear Gaussian three-factor model. Our new model adds realistic time variation to the quantity of risk while inheriting the flexibility of the standard Gaussian model to explain the time-variation in the price of risk. Hence, the model is able to jointly capture the time-variation in expected returns and volatilities of bond returns. Moreover, our nonlinear model reveals a sharp increase of the mean and volatility of expected excess returns during the early eighties, something that is missed by affine models.

We explore the implications of the new model for investments in U.S. Treasury bonds. The nonlinear model predicts higher Sharpe ratios than the standard model in low-volatility periods. Intuitively, volatility is equal to the sample average in the standard model and thus this model overshoots volatility and underpredicts Sharpe ratios in low-volatility periods. In high-volatility periods, particularly during the early 80s, variations in expected excess returns in the nonlinear model are mostly driven by

[1]See Dai and Singleton (2003) and Duffee (2010b) and the references therein.

[2]Examples include Sangvinatsos and Wachter (2005), Cochrane and Piazzesi (2008), Duffee (2010b), and Joslin, Priebsch, and Singleton (2012).

variations in volatility. Thus, the high expected excess returns during these periods are primarily driven by a higher quantity of risk rather than a higher compensation for risk and Sharpe ratios in the two models are similar.

There is strong evidence that an economically significant part of excess returns is not explained by linear combinations of yields.[3] This finding motivates Gaussian "hidden factor" models where one or more factors are orthogonal to yields but help explain expected excess returns. Our paper highlights an alternative channel: a nonlinear relation between expected excess returns and yields. Specifically, we regress expected excess returns implied by the nonlinear model onto yields and show that the error of this regression accounts for 14% to 29% of the total variance depending on forecast horizon and bond maturity. Duffee (2011a) and Joslin, Priebsch, and Singleton (2012) generate similar results using five factors whereas three factors are sufficient in our nonlinear model.

The standard procedure in the term structure literature is to specify the short rate and the market prices of risk as functions of the state variables. Instead, we model the functional form of the stochastic discount factor directly by multiplying the stochastic discount factor from a Gaussian term structure model with the term $1 + \gamma e^{-\beta X}$ where β and γ are parameters and X is the Gaussian state vector. This functional form is a special case of the stochastic discount factor that arises in many equilibrium models in the literature as we show in Appendix B. In such models the stochastic discount factor can be decomposed into a weighted average of different representative agent models. Importantly, the weights on the different models are time-varying and this is a source of time-varying risk premia and volatility of bond returns.

[3]Recent papers on this topic include Ludvigson and Ng (2009), Cooper and Priestley (2009), Cieslak and Povala (2010), Duffee (2011a), Joslin, Priebsch, and Singleton (2012), and Chernov and Mueller (2012).

The nonlinear and standard Gaussian model lead to similar predictions for expected excess returns except for the early 80s where they are higher in the nonlinear model. To distinguish between the two models' return implications, we regress realized excess returns on model-implied expected excess return. The slope coefficient should be close to one. We find a slope coefficient between 0.95 and 1.15 for the nonlinear model and higher than one for the Gaussian model (1.05 to 1.77) supporting the implications of the nonlinear model. Interestingly, our results on expected excess returns are similar to the findings of Dai, Singleton, and Yang (2007) who use a regime-switching Gaussian three-factor model. While the Gaussian model is a special case of both models our nonlinear model only increases the number of parameters from 23 to 27 whereas the regime-switching model in Dai, Singleton, and Yang (2007) has 56 parameters.

To study the second moments of yields we look at the variation in the one-month ahead conditional volatility of yields. Conditional volatilities in the nonlinear model show high correlation with GARCH estimates. In contrast to affine models with stochastic volatility, the nonlinear model captures the spike in volatilities during the 80s: the highest one-month conditional yield volatility for a 2-year bond during the sample is 161 basis points according to GARCH, 133 basis points according to the nonlinear model, while it is less than 70 basis points in affine models with stochastic volatility (see Jacobs and Karoui (2009)). In the Gaussian model volatility is constant at 41 basis points.

Our paper is not the first to propose a term structure model outside the class of Gaussian models. General affine models (Duffie and Kan (1996) and Dai and Singleton (2000)) allow for stochastic volatility, but in contrast to our nonlinear model cannot match both stochastic volatility and time-variation in expected excess returns (see for example Duffee (2002) and Dai and Singleton (2002)). Quadratic term structure models have been proposed by among others Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003). Cheng and Scaillet (2007) show that quadratic term structure models can be embedded into the affine class using an augmented state vector, suggesting that the non-linearity in this class of models is of restricted nature. Consistent with this, Ahn, Dittmar, and Gallant (2002) find that quadratic term structure models are not able to generate the level of conditional volatility observed for short- and intermediate-term bond yields. We show that the nonlinear model can match the time variation of conditional volatility for both short and long maturity bonds. Ahn, Dittmar, Gallant, and Gao (2003) propose a class of nonlinear term structure models based on the inverted square-root model of Ahn and Gao (1999), but in contrast to our nonlinear model they do not provide closed-form solutions for bond prices.

The rest of the paper is organized as follows. Section II describes the model. Section III estimates the model and Section IV presents the empirical results. Section V concludes.

II The Model

In this section we present a model of the term structure of interest rates.[4] Uncertainty is represented by a d-dimensional Brownian motion $W(t) = (W_1(t), ..., W_d(t))'$. There is a d-dimensional Gaussian state vector $X(t)$ that follows the dynamics

$$dX(t) = \kappa \left(\bar{X} - X(t)\right) \, dt + \Sigma \, dW(t), \tag{1}$$

where $\bar{X}$ is d–dimensional and κ and Σ are $d \times d$–dimensional.

[4]We introduce a more general class of nonlinear term structure models in Appendix A.

II.A The Stochastic Discount Factor

We assume that there is no arbitrage and hence there exists a strictly positive state price density or stochastic discount factor $M(t)$. Let γ denote a nonnegative constant, β a d-dimensional vector, and $M_0(t)$ a strictly positive stochastic process. The stochastic discount factor is defined as

$$M(t) = M_0(t) \left(1 + \gamma e^{-\beta' X(t)}\right). \tag{2}$$

Equation (2) is a key departure from standard term structure models (Vasicek (1977), Cox, Ingersoll, and Ross (1985), Duffie and Kan (1996), and Dai and Singleton (2000)). Rather than specifying the short rate and the market price of risk, which in turn pins down the state price density, we specify the functional form of the state price density directly.[5]

To keep the model comparable to the existing literature on Gaussian term structure models we introduce a base model for which $M_0(t)$ is the stochastic discount factor. The dynamics of $M_0(t)$ are

$$\frac{dM_0(t)}{M_0(t)} = -r_0(t)dt - \Lambda_0(t)' dW(t), \tag{3}$$

where $r_0(t)$ and $\Lambda_0(t)$ are affine functions of the state vector $X(t)$. Specifically,

$$r_0(X) = \rho_{0,0} + \rho_{0,X}' X, \tag{4}$$

$$\Lambda_0(X) = \lambda_{0,0} + \lambda_{0,X} X, \tag{5}$$

where $\rho_{0,0}$ is a scalar, $\rho_{0,X}$ and $\lambda_{0,0}$ are d-dimensional vectors, and $\lambda_{0,X}$ is a $d \times d$-dimensional matrix. It is well known that bond prices in the base model belong to the

[5] Constantinides (1992) and Rogers (1997) also specify the functional form of the state price density directly and provide closed form solutions for bond prices.

class of essentially affine term structure models (Duffee (2002) and Dai and Singleton (2002)). We now provide closed form solutions for bond prices in the general model.

II.B Closed-Form Bond Prices

Let $P^{(\tau)}(t)$ denote the price at time t of a zero-coupon bond that matures in τ years. Specifically,

$$P^{(\tau)}(t) = \mathrm{E}_t\left[\frac{M(t+\tau)}{M(t)}\right]. \tag{6}$$

We show in the next theorem that the price of a bond is a weighted average of bond prices in artificial economies that belong to the class of essentially affine Gaussian term structure models.

Theorem 1. *The price of a zero-coupon bond that matures in τ years is*

$$P^{(\tau)}(t) = s(t)P_0^{(\tau)}(t) + (1-s(t))P_1^{(\tau)}(t), \tag{7}$$

where

$$s(t) = \frac{1}{1+\gamma e^{-\beta' X(t)}} \tag{8}$$

$$P_n^{(\tau)}(t) = e^{A_n^*(\tau) - B_n^*(\tau)' X(t)}. \tag{9}$$

The coefficient $A_n^*(\tau)$ and the d-dimensional vector $B_n^*(\tau)$ solve the ordinary differential equations

$$\frac{dA_n^*(\tau)}{d\tau} = \frac{1}{2}B_n^*(\tau)'\Sigma\Sigma' B_n^*(\tau) - B_n^*(\tau)'\left(\kappa\bar{X} - \Sigma\lambda_{n,0}\right) - \rho_{n,0}, \quad A_n^*(0) = 0, \tag{10}$$

$$\frac{dB_n^*(\tau)}{d\tau} = -\left(\kappa + \Sigma\lambda_{n,X}\right)' B_n^*(\tau) + \rho_{n,X}, \quad B_n^*(0) = 0_d, \tag{11}$$

where

$$\rho_{n,0} = \rho_{0,0} + n\beta'\kappa\bar{X} - n\beta'\Sigma\lambda_{0,0} - \frac{1}{2}n^2\beta'\Sigma\Sigma'\beta, \tag{12}$$

$$\rho_{n,X} = \rho_{0,X} - n\kappa'\beta - n\lambda_{0,X}'\Sigma'\beta, \tag{13}$$

$$\lambda_{n,0} = \lambda_{0,0} + n\Sigma'\beta, \tag{14}$$

$$\lambda_{n,X} = \lambda_{0,X}. \tag{15}$$

The proof of this theorem is given in Appendix A. To provide some intuition we rewrite the bond pricing equation (6) and use the fact that $s(t) = M_0(t)/M(t)$:

$$P^{(\tau)}(t) = s(t)E_t\left[\frac{M_0(T)}{M_0(t)}\right] + (1 - s(t))E_t\left[\frac{\gamma e^{-\beta' X(T)}M_0(T)}{\gamma e^{-\beta' X(t)}M_0(t)}\right] \tag{16}$$

Both expectations are equal to bond prices in artificial economies with discount factors $M_0(t)$ and $M_1(t) = \gamma e^{-\beta' X(t)}M_0(t)$, respectively. These bond prices belong to the class of essentially affine term structure models and hence $P^{(\tau)}(t)$ can be computed in closed form.

II.C The Short Rate and the Price of Risk

Applying Ito's lemma to equation (2) leads to the dynamics of the stochastic discount factor:

$$\frac{dM(t)}{M(t)} = -r(t)\,dt - \Lambda(t)'dW(t), \tag{17}$$

where both the short rate $r(t)$ and the market price of risk $\Lambda(t)$ are nonlinear functions of the state vector $X(t)$ given in equations (18) and (19), respectively. The short rate is given by

$$r(t) = r_0(t) + (1 - s(t))\left(\beta'\kappa\left(\bar{X} - X(t)\right) - \beta'\Sigma\Lambda_0(t) - \frac{1}{2}\beta'\Sigma\Sigma'\beta\right). \tag{18}$$

Our model allows the short rate to be nonlinear in the state variables without losing the tractability of closed form solutions of bond prices and a Gaussian state space.[6]

The d-dimensional market price of risk is given by

$$\Lambda(t) = \Lambda_0(t) + (1 - s(t))\,\Sigma'\beta. \tag{19}$$

From equation (19) we can see that even if the market prices of risk in the base model are constant, the market prices of risks in the general model are stochastic due to variations in the weight $s(t)$. When $s(t)$ approaches zero or one, then $\Lambda(t)$ approaches the market price of risk of an essentially affine Gaussian model.

II.D Expected Return, Volatility, and Sharpe Ratio

We know that the bond price is a weighted average of exponential affine bond prices (see equation (7)). Hence, variations of instantaneous bond returns are due to variations in the two artificial bond prices $P_0^{(\tau)}(t)$ and $P_1^{(\tau)}(t)$ and due to variations in the weight $s(t)$. Specifically, the dynamics of the bond price $P^{(\tau)}(t)$ are

$$\frac{dP^{(\tau)}(t)}{P^{(\tau)}(t)} = \left(r(t) + e^{(\tau)}(t)\right)\, dt + v^{(\tau)}(t)'\, dW(t), \tag{20}$$

where $e^{(\tau)}(t)$ denotes the instantaneous expected excess return and $v^{(\tau)}(t)$ denotes the local volatility vector of a zero-coupon bond that matures in τ years.

The local volatility of the bond is given by

$$v^{(\tau)}(t) = -\Sigma'\left(\omega^{(\tau)}(t)B_0^*(\tau) + \left(1 - \omega^{(\tau)}(t)\right)B_1^*(\tau) + \beta\left(s(t) - \omega^{(\tau)}(t)\right)\right), \tag{21}$$

[6]Chan, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996a), Ait-Sahalia (1996b), Stanton (1997), Pritsker (1998), Chapman and Pearson (2000), Ang and Bekaert (2002), and Jones (2003) study the nonlinearity of the short rate. Jermann (2013) and Richard (2013) study nonlinear term structure models but they do not get closed form solutions for bond prices.

where $\omega^{(\tau)}(t)$ denotes the contribution of $P_0^{(\tau)}(t)$ to the bond price $P^{(\tau)}(t)$:

$$\omega^{(\tau)}(t) = \frac{P_0^{(\tau)}(t)s(t)}{P^{(\tau)}(t)}. \tag{22}$$

When $s(t)$ approaches zero or one, then $v^{(\tau)}(t)$ approaches the constant local volatility of a Gaussian model. However, in contrast to the market price of risk, the local volatility is not bounded by the local volatilities of the two Gaussian models.

The instantaneous expected excess return of the bond is

$$e^{(\tau)}(t) = \Lambda(t)' v^{(\tau)}(t). \tag{23}$$

We can see from equations (17)-(23) that our nonlinear term structure model differs from the essentially affine Gaussian base model in two important aspects. First, the volatilities of bond returns and yields are time-varying and hence expected excess returns are moving with the price and the quantity of risk.[7] Second, the short rate $r(t)$, the instantaneous volatility $v^{(\tau)}(t)$, and the instantaneous expected excess return $e^{(\tau)}(t)$ are nonlinear functions of $X(t)$.

III Estimation

In this section, we estimate a standard and a nonlinear three-factor essentially affine Gaussian model. The nonlinear model has the same number of factors and the number of parameter increases from 23 to 27.

[7] The instantaneous volatility of the bond yield is $-\frac{1}{\tau}v^{(\tau)}(t)$.

III.A Data

The models are estimated using a monthly panel of zero-coupon Treasury bond yields. We use month-end (continuously compounded) 1-, 2-, 3-, 4-, and 5-years zero-coupon yields extracted from U.S. Treasury security prices by the method of Fama and Bliss (1987). The data is from the Center for Research in Security Prices and covers the period 1952:6 to 2011:12.

III.B Estimation Methodology

We use the Kalman filter to estimate the standard Gaussian model and the unscented Kalman filter to estimate the nonlinear Gaussian model. Christoffersen, Dorion, Jacobs, and Karoui (2012) find that the unscented Kalman filter works well in estimating affine term structure models when highly nonlinear instruments are observed. We briefly discuss the setup but refer to Carr and Wu (2009) and Schwartz and Trolle (2012) for a detailed description of this filter.

We stack the N observed yields in month t in the vector $y(t)$, and set the model up in state-space form. The measurement equation is

$$y(t) = f(X(t)) + \epsilon(t), \qquad \epsilon(t) \sim N(0, \sigma I_N), \tag{24}$$

where $f(\cdot)$ is the function determining the relation between the latent variables and yields.[8] We use the Kalman filter if $f(\cdot)$ is linear and the unscented Kalman filter if $f(\cdot)$ is nonlinear. The transition equation for the latent variables is

$$X(t+1) = C + DX(t) + \eta(t+1), \qquad \eta(t) \sim N(0, Q), \tag{25}$$

[8] $f = (f_1, ..., f_5)'$ where $f_i(X(t)) = -\frac{1}{\tau_i} \ln\left(P^{(\tau_i)}(X(t))\right)$ with $P^{(\tau)}(X(t))$ given in equation (6) and $\tau_i = i$.

where C is a vector and D is a matrix that enters the one-month ahead expectation of $X(t)$; i.e. $E_t(X(t+1)) = C + DX(t)$. Q is the covariance matrix of $X(t+1)$ given $X(t)$ which is constant since $X(t)$ is Gaussian.

Since not all of the parameters are identified, we apply the normalizations proposed in Dai and Singleton (2000).[9] Specifically, for the dynamics of X given in equation (1) we assume that the mean reversion matrix, κ, is lower triangular, the mean of the state variables, $\overline{X}$, is the zero vector, and that the local volatility, σ_X, is the identity matrix. Recent literature (Collin-Dufresne, Goldstein, and Jones (2008), Joslin, Singleton, and Zhu (2011), Christensen, Diebold, and Rudebusch (2011), Hamilton and Wu (2012), and Joslin, Le, and Singleton (2013)) propose other parameterizations to ease the estimation of affine term structure models. Since these results cannot be applied to the nonlinear model, we choose the Dai and Singleton (2000) normalizations.

III.C Estimation Results

Parameter estimates and log-likelihood values are reported in Table 1. The nonlinear model has four additional parameters (the scalar γ and the three-dimensional vector β) and a log-likelihood value that is 258.3 higher than the Gaussian model. The statistical significance of the nonlinear Gaussian model cannot be tested using a standard Likelihood Ratio (LR) test because the parameter vector β is not identified under the null hypothesis of having a Gaussian model ($\gamma = 0$) and hence the LR is not asymptotically χ^2-distributed. Garcia (1998) derives the distribution of the LR statistic for a two-state Markov model and finds the 5% critical value to be 14.11 while the critical value in the χ^2-distribution is 9.5. In our case the LR statistic of 258.3 is well in excess of the 5% critical value of 9.5 in the χ^2-distribution with

[9] As is often the case in multi-factor term structure models, individual parameters are not easily interpretable and in the following we focus on economic implications of the model.

four degrees of freedom, suggesting that the nonlinear extension is statistically highly significant. The standard deviation of pricing errors, σ, is almost the same in the two models. This implies that the significance of the nonlinear model does not come from an improved cross-sectional fit of yields.

The bond price is a weighted average of two Gaussian bond prices (see Theorem 1). If the weight $s(t)$ is 1 the bond price in the nonlinear model collapses to the bond price in the Gaussian base model.[10] Figure 1 shows the weight on the base model. We see that in the 70's and 80's the weight is significantly below one, so nonlinearity becomes particularly important during these periods. The shaded areas in the figure are the NBER recessions and we see that the weight moves away from one during recessions.

IV Empirical Results

We focus in this section on the empirical properties of the nonlinear Gaussian term structure model and compare it to the standard Gaussian model.

IV.A Expected Excess Returns

Expected excess returns of U.S. Treasury bonds vary over time as documented in Fama (1984), Fama and Bliss (1987) and Campbell and Shiller (1991) (CS). CS document this by regressing future yield changes on the scaled slope of the yield curve. The slope regression coefficient is one if excess holding period returns are constant, but

[10]The representation of the bond price as weighted average of two different bond prices is similar to a regime switching model with unobservable regimes. In this case $s(X(t))$ would denote the probability of being in regime 0 conditional on the state of the economy $X(t)$. However, it is not clear that there exist Markov-transition probabilities such that the conditional distribution of yields can be written as a weighted average of the conditional distribution of yields in regime 0 and 1 and thus be consistent with Bayes rule.

CS find negative regression coefficients. It is well documented that Gaussian models can capture the predictability of excess bond returns as measured through the CS regression coefficients while affine models with stochastic volatility cannot.[11] Panel A in Table 2 shows that both the Gaussian and nonlinear Gaussian model match the CS regression coefficients quite well; i.e. the coefficients are negative and decreasing in the maturity of the bond consistent with the evidence in the data.

Figure 2 shows expected one-year log excess returns implied by the Gaussian and nonlinear Gaussian model.[12] The models have similar predictions for excess returns apart from the time of the monetary experiment in the early eighties. In this period the nonlinear model predicts higher excess returns than in any other time in the sample period. Moreover, they are twice as high as in the Gaussian model. Consistent with the predictions of the nonlinear model the early eighties had two instances where realized excess returns were higher than in any other time in the sample period. For instance, the excess return of the five year bond exceeded 16% twice during that period.

To formally test the ability of the models to capture expected excess returns, we run a regression of realized excess returns on expected excess returns. The results are reported in Panel B of Table 2. If the model captures expected excess returns well, then the slope coefficient should be one and the constant zero in all regressions. We see that the slope coefficients are close to one in the nonlinear model while they are too high for the Gaussian model. This holds true for all bond maturities and holding horizons. For example, the average slope coefficient for a holding period of one year is 1.05 in the nonlinear model while it is 1.65 in the Gaussian model. Furthermore

[11] See Dai and Singleton (2002), Tang and Xia (2007), and Feldhütter (2008). Almeida, Graveline, and Joslin (2011) show that if options are included in the estimation, affine models with stochastic volatility can match the CS regression coefficients.

[12] Moments of yields and returns are easily calculated using Termite-Gauss quadrature, see Appendix C for details. In the rest of the paper we use Hermite-Gauss quadrature when we do not have closed-form solutions for expectations or variances.

the constant α is closer to zero in the nonlinear model than in the Gaussian model.

To check the ability of a model's expected excess return to explain realized excess returns relative to the expectation hypothesis with constant expected excess returns we define the measure "fraction of variance explained" as:

$$\text{FVE} = 1 - \frac{\frac{1}{T}\sum_{t=1}^{T}(rx_{t,t+n}^{\tau} - E_t(rx_{t,t+n}^{\tau}))^2}{\frac{1}{T}\sum_{t=1}^{T}(rx_{t,t+n}^{\tau} - \overline{rx^{\tau}})^2}, \tag{26}$$

where $rx_{t,t+n}^{\tau}$ is the n-year return on a bond with maturity τ in excess of the n-year return on a bond with maturity n ($\tau > n$). The R^2 of the regression of realized on expected excess returns is equal to FVE if the slope is one and the constant is zero, otherwise the R^2 is an upper bound for FVE. Panel B of Table 2 shows the coefficients of this regression, the R^2, and the FVE for different bond maturities and holding periods.[13] For one-year horizons FVE's in both the nonlinear and Gaussian model are similar. The reason is that although the nonlinear model captures high excess returns in the early eighties, volatility is also high leading to occasional high negative excess returns and noise. This noise created by volatility is attenuated when looking at longer holding horizons in which case the FVE's in the nonlinear model increases to more than double of those in the Gaussian model.

Taken together, the nonlinear model captures a rise in expected excess returns in the early eighties missed by the Gaussian model. This finding is important not only for term structure modeling but also for the common approach of predicting excess returns with vector-autoregressive (VAR) models (classic examples are Campbell and Ammer (1993) and Ang and Piazzesi (2003)). Joslin, Singleton, and Zhu (2011) show that conditional expected excess returns in a Gaussian model without parameter restrictions are identical to those from an unrestricted VAR model. Thus, any three-

[13]Because we look at returns over several years in Panel B, the difference between log excess returns and raw excess returns increases and we therefore report the regression for raw returns. The table for log excess returns shows similar results for the regression coefficients, but the average FVE is 7.6% instead of 12.5% for the nonlinear model and 5.5% instead of 9.0% for the Gaussian model.

factor VAR with yields as factors would miss the rise in expected excess returns.

IV.B Stochastic Volatilities

In this section we study how well volatilities of yields (and thus volatilities of excess returns) in the data are matched by the nonlinear model. Yield volatilities in the Gaussian model are constant over time and this is a major shortcoming when studying investment opportunities in the bond market (see Section IV.C) or when studying pricing and risk management of fixed income securities.

Volatility of yields varies over time and is persistent (Singleton (2006) and the references therein). To show that our nonlinear model is able to capture the time variation in yield volatilities we follow the literature and compare model-implied conditional volatility of monthly changes in yields for all five bonds with estimates of an EGARCH(1,1) model (see for example Jacobs and Karoui (2009), Almeida, Graveline, and Joslin (2011), and Kim and Singleton (2011)). Panel A of Table 3 shows that the correlations between model-implied (calculated using Gaussian quadrature) and actual (measured with the EGARCH model) volatilities are quite high, between 71% and 75%. These correlations are similar in magnitude to the correlations found in affine models with one or more CIR processes (Feldhütter (2008) and Jacobs and Karoui (2009)).

We follow the approach of Jacobs and Karoui (2009) and Kim and Singleton (2011) and regress EGARCH volatility on model-implied volatility. Panel A shows the coefficients. Although widely used, EGARCH volatility is a model-dependent estimate of volatility and it is not clear that the slope coefficients should be one. We therefore simulate 100,000 months of yields from the nonlinear model, use the simulated yields to estimate EGARCH volatility, and regress estimated EGARCH volatility on model-implied volatility in the 100,000 months. The resulting regression coefficients are $\overline{\alpha}$ and $\overline{\beta}$ in Panel A. We see that the true simulated slope coefficients are 0.65-0.71 and that the actual slope coefficients are in the range 0.71-0.87 and statistically insignificant from the simulated regression coefficients. This shows that the nonlinear model captures well the dynamics of volatility as measured through EGARCH. In contrast the slope coefficients of the essentially affine three-factor model with one stochastic volatility factor reported in Table 5 of Jacobs and Karoui (2009) are 1.92, 1.49, and 1.13 for the 1-, 2-, and 5-year yields.[14] This illustrates the difficulty affine models with stochastic volatility have in matching volatility; i.e. they tend to have too little variation in volatility and this is more pronounced for short maturities.

We can also see how the nonlinear model better than affine models captures spikes in volatility by looking at the right tail of the distribution of conditional volatilities in Panel B of Table 3. The most volatile period in the sample is the early 80s and the 95th and 99th quantiles reflect the high volatility during this period. The quantiles for the conditional yield of the 1-year bond in the data according to the EGARCH model are 96bps and 164bps and they are fairly well matched by 100bps and 138bps in the nonlinear model. For standard affine models with one stochastic volatility factor the maximal conditional volatility is around 70bps in the early 80s (Jacobs and Karoui (2009) Fig. 1-3 Panel A). In the regime-switching model of Dai, Singleton, and Yang (2007) actual volatility is between two and three times higher than model-implied volatility in the early 80s.[15] Overall, the evidence suggests that the nonlinear model captures the high volatility in the early 80s better than existing term structure models.

Figure 3 shows the volatility of excess returns for a two, three, four, and five year bond. The graphs show that the volatility implied by the nonlinear model is

[14]To check what the regression coefficients in the essentially affine model should be we simulated yields from the model using the parameter estimates in Jacobs and Karoui (2009) and found for the 1-, 2-, and 5-year yield $\overline{\alpha}$ to be 0.00036, 0.00036, and 0.00031 and $\overline{\beta}$ to be 0.78, 0.77, and 0.77.

[15]This is for the conditional volatility of the 10-year yield in their Figure 8. They do not report results for bonds with shorter maturity.

lower in the first and last third of the sample than predicted by the Gaussian model. Moreover, the nonlinear model captures the volatility increases during recessions and the volatility spike during the early eighties.

IV.C Sharpe Ratios

We measure the investment opportunity of a bond at time t with its Sharpe ratio

$$S_t = \frac{E_t(rx^{(\tau)}_{t,t+12})}{\sqrt{Var_t(rx^{(\tau)}_{t,t+12})}}, \tag{27}$$

where $rx^{(\tau)}_{t,t+12}$ is the one-year log return on a bond with maturity τ in excess of the one-year log return of a one-year bond. Gaussian models are benchmark models when examining the time-varying investment opportunities in the bond market. Viewing the risk-return tradeoff in the bond market through the lens of a Gaussian model can lead to inaccurate conclusions because bond return volatilities in Gaussian models are constant and hence all the variation in Sharpe ratios must come from the variation in expected excess returns. As Duffee (2010a) points out: "existing dynamic term structure models are insufficiently flexible to capture the empirical dynamics of both conditional means and conditional volatilities. Thus either the numerator or denominator of the conditional Sharpe ratio is likely misspecified." Since the nonlinear model captures the time variation in both moments well, the model provides more realistic estimates of conditional Sharpe ratios than the Gaussian benchmark model.

Figure 4 shows conditional Sharpe ratios in the Gaussian and the nonlinear model. There is a substantial difference between the Sharpe ratios of the Gaussian and nonlinear model in the calm periods of the data sample. In these periods the two models agree on expected excess returns as shown in Section IV.A, but volatility in the Gaussian model is too high because the volatility has to match average sample volatility.

As a consequence, the Gaussian model predicts Sharpe ratios that are too low. For the nonlinear model most of the variation in expected excess returns in the calm periods comes from variation in the Sharpe ratio. Why are Sharpe ratios similar in the 80s? In the 80s the weight s in equation (7) is dropping far below one and the nonlinear relation between yields and the Gaussian state variables becomes stronger. Market prices of risk do not change significantly, but yield volatilities spike up. Variations in expected excess returns in this period are mostly driven by variations in volatility. Thus, the high expected returns during the eighties are primarily a result of a higher quantity of risk and not a higher compensation for risk.

IV.D Hidden Information

Standard affine term structure models imply that expected excess returns are fully explained by a linear combination of yields. However, there is strong evidence that an economically significant part of expected excess returns are not explained by linear combinations of yields.[16]

This finding motivates Gaussian "hidden factor" models where one or more factors determine excess returns but are partially unrelated to linear combinations of yields. Hidden factors show up either through explicit parameter restrictions as in Joslin, Priebsch, and Singleton (2012) or through filtering in a term structure model with at least five factors as in Duffee (2011a).

Our nonlinear model highlights an alternative channel through which excess returns are imperfectly correlated with yields. In the model expected excess returns are nonlinearly related to yields and therefore a part of expected excess returns is "hidden" from a linear combination of yields. Is the nonlinearity strong enough to be

[16]Recent papers in this topic include Ludvigson and Ng (2009), Cooper and Priestley (2009), Cieslak and Povala (2010), Duffee (2011a), Joslin, Priebsch, and Singleton (2012), and Chernov and Mueller (2012).

empirically relevant? To answer this question we follow Duffee (2011a) and calculate the ratio of the variance of expected log excess returns projected onto model-implied yields divided by the variance of the true expected log excess returns.[17] In the Gaussian model this ratio is 1. Table 4 shows that in the nonlinear model this ratio is between 0.71 and 0.86 depending on the forecast horizon and the bond maturity. Compared with the range 0.53 to 0.70 reported in Duffee (2011a)'s five-factor model, this suggests that nonlinearities are important in understanding the unspanned part of excess returns.

Is the linearly hidden part of excess returns related to macro variables? To answer this we regress expected excess returns on yields and call the residual the linearly hidden part of expected excess returns. Then, we regress this hidden part on inflation and industrial production growth. Table 5 shows that there is a significant negative relation between the linearly hidden part and inflation. This relation occurs because inflation is correlated with the amount of nonlinearity in the nonlinear model, not because inflation has predictive power above what is contained in the yield curve. The amount that inflation and industrial production growth explain of the linearly hidden part (R^2 of 6%) is similar to the amount these macro variables explain of Duffee (2011a)'s hidden factor (R^2 of 8%).[18] The conclusion is that in order to better understand how much predictive power - beyond what is contained in the yield curve - macro variables have in explaining expected excess, it is important to control for nonlinearity.

Linearly hidden factors and "truly" hidden factors in the spirit of Duffee (2011a) and Joslin, Priebsch, and Singleton (2012) are likely to both play an important rule in understanding expected excess returns. Consistent with this view, Table 5 shows that Duffee's hidden factor is uncorrelated with the linearly hidden part of excess returns

[17]We regress expected log excess returns on model-implied yields instead of actual yields to highlight the contribution of the nonlinearity.

[18]See Table 6 in Duffee (2011a)

in the nonlinear model. However, from a modelling perspective the linearly hidden part appears more naturally than the "truly" hidden part. As just shown a significant linearly hidden factor appears in a nonlinear three-factor model. In contrast, Duffee (2011a) shows that at least five latent factors along with yield measurement errors are needed to generate a partially hidden factor. Unless economically motivated restrictions are imposed on parameters, a five-factor model leads to huge Sharpe ratios due to overfitting (Duffee (2010a)) .

IV.E Cross-sectional fit of three-factor models

The nonlinear description of yield dynamics in the nonlinear model allows us to capture the time variation in the mean and volatility of excess bond returns. Empirically, Balduzzi and Chiang (2012) show that in the cross-section there is an almost linear relation between yields. To see if the nonlinear model captures the cross-sectional linearity we follow Duffee (2011b) and determine the principal components of zero-coupon bond yields with maturities ranging from one to five years and regress the yield of each bond on all five principal components. The results for the data (715 monthly observations) and the two models are shown in Table 6. The results for the models are based on one million simulated observations.

As is well known we see that the first three principal components describe almost all the variation of bond yields in the data. Panel A of Table 6 shows that they also describe almost all the variation of bond yields in the nonlinear Gaussian model. Moreover, Panel B of Table 6 shows that the loading for each yield on the level, slope, and curvature factor are very similar to the data (and the Gaussian model). Interestingly, although the explanatory power of the fourth and fifth principal component are low, the loadings on these factors in the nonlinear model are similar to the loadings in the data. In contrast, the first three principal components describe by construction all the variation of bond yields in the Gaussian model and hence the loadings for all yields on the fourth and fifth principal component are zero.

To conclude, the cross-sectional variation of bond yields implied by the nonlinear model is well explained by the first three principal components and no yield breaks this linear relation.

V Conclusion

We introduce a new reduced form term structure model where the short rate and market prices of risk are nonlinear functions of Gaussian state variables but yields are nevertheless given in closed form. We estimate the model on a long time-series of U.S. Treasury yields that includes the monetary experiment of the early eighties and show that the model captures the time variation in expected returns and volatilities of Treasury bonds well. We also show that during low volatility regimes Treasury bonds are more attractive investment than standard Gaussian models predict and that nonlinearities can help explain why expected excess returns are not explained by a linear combination of yields.

Our nonlinear model successfully captures the spike of expected excess returns and volatilities during the extreme period of the early eighties while preserving the linear relation in the cross section of yields.[19] This period is particularly challenging for affine term structure models to explain. We are currently working on including options in the estimation to study asset pricing implications during the post Volcker-period.

Although our empirical analysis has focused on a nonlinear generalization of an affine Gaussian model, it is possible to generalize a wide range of term structure mod-

[19]This suggests that our model can be used to measure and price disaster risk (e.g. Barro (2006)).

els such as affine models with stochastic volatility, quadratic models, and nonlinear models. Our generalization introduces new dynamics for bond returns while keeping the new model as tractable as the standard model. Furthermore, the method extends to processes such as jump-diffusions and continuous time Markov chains. We explore this in Feldhütter, Heyerdahl-Larsen, and Illeditsch (2013).

References

Ahn, D.-H, R.F. Dittmar, and A.R. Gallant, 2002, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* 15, 243–288.

Ahn, Dong-Hyun, Robert F. Dittmar, A. Ronald Gallant, and Bin Gao, 2003, Purebred or hybrid? Reproducing the volatility in term structure dynamics, *Journal of Econometrics* 116, 147–180.

Ahn, Dong-Hyun, and Bin Gao, 1999, A parametric non-linear model of term structure dynamics, *Review of Financial Studies* 12, 721–762.

Ait-Sahalia, Yacine, 1996a, Nonparametric pricing of interest rate derivative securities, *Econometrica* 64, 527–560.

——— , 1996b, Testing continous-time models of the spot interest rate, *Review of Financial Studies* 9, 385–426.

Almeida, Caio, Jeremy J. Graveline, and Scott Joslin, 2011, Do interest rate options contain information about excess bond returns?, *Journal of Econometrics* 164, 35–44.

Ang, Andrew, and Geert Bekaert, 2002, Short rate nonlinearities and regime switches, *Journal of Economic Dynamics and Control* 26, 1243–1274.

Ang, Andrew, and Monica Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50, 745–787.

Balduzzi, Pierluigi, and I-Hsuan Ethan Chiang, 2012, A simple test of the affine class of term structure models, *Review of Asset Pricing Studies* 2, 203–244.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *The Quarterly Journal of Economics* 121, 823–866.

Basak, Suleyman, 2000, A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk, *Journal of Economic Dynamics and Control* 24, 63–95.

——— , and Domenico Cuoco, 1998, An equilibrium model with restricted stock market participation, *Review of Financial Studies* 11, 309–341.

Bhamra, Harjoat, and Raman Uppal, 2010, Asset prices with heterogeneity in preferences and beliefs, University of British Columbia.

Campbell, John Y., and John Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance* 48, 3–37.

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

Campbell, John Y., and Rober J. Shiller, 1991, Yield spread and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Carr, Peter, and Liuren Wu, 2009, Stock options and credit default swaps: A joint framework for valuation amd estimation, *Journal of Financial Econometrics* pp. 1–41.

Chan, K.C., Andrew Karolyi, Francis Longstaff, and Anthony B. Sanders, 1992, An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance* 47, 1209–1227.

Chapman, David A., and Neil D. Pearson, 2000, Is the short rate drift actually nonlinear?, *Journal of Finance* LV, 355–388.

Chen, Hui, and Scott Joslin, 2012, Generalized transform analysis of affine processes and applications in finance, *Review of Financial Studies* 25, 2225–2256.

Cheng, Peng, and Olivier Scaillet, 2007, Linear-quadratic jump-diffusion modelling, *Mathematical Finance* 17, 575–598.

Chernov, Mikhail, and Philippe Mueller, 2012, The term structure of inflation expectations, *Journal of Financial Economics* 106, 367–394.

Christensen, Jens H.E., Francis X. Diebold, and Glenn D. Rudebusch, 2011, The affine arbitrage-free class of nelson-siegel term structure models, *Journal of Econometrics* 164, 4–12.

Christoffersen, Peter, Christian Dorion, Kris Jacobs, and Lotfi Karoui, 2012, Nonlinear Kalman Filtering in Affine Term Structure Models, *Working Paper*.

Cieslak, Anna, and Pavel Povala, 2010, Understanding bond risk premia, Working paper.

Cochrane, John, and Monika Piazzesi, 2008, Decomposing the yield curve, Working paper.

Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara, 2008, Two trees, *Review of Financial Studies* 21, 347–385.

Collin-Dufresne, Pierre, Robert S. Goldstein, and Christopher S. Jones, 2008, Identification of Maximal Affine Term Structure Models, *Journal of Finance* 63, 743–795.

Constantinides, George M., 1992, A theory of the nominal term structure of interest rates, *Review of Financial Studies* 5, 531–552.

Cooper, Ilan, and Richard Priestley, 2009, Time-Varying Risk Premiums and the Output Gap, *Review of Financial Studies* 22, 2601–2633.

Cox, John. C., Jonathan E. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, p. 385–408.

Cuoco, D., and H. He, 1994, Dynamic equilibrium in infinite-dimensional economies with incomplete financial markets, Wharton School, University of Pennsylvania.

Dai, Q., and K. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

———, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631–678.

Dai, Qiang, and Kenneth J. Singleton, 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–441.

———, and Wei Yang, 2007, Regime shifts in a dynamic term structure model of u.s. treasury bond yields, *Review of Financial Studies* 20, 1669–1706.

Duffee, G., 2002, Term premia and interest rate forecast in affine models, *Journal of Finance* 57, 405–443.

Duffee, Gregory, 2010a, Sharpe ratios in term structure models, Working paper.

Duffee, Greg, 2011a, Information in (and not in) the term structure, *Review of Financial Studies* 24, 2895–2934.

Duffee, Gregory R., 2010b, Sharpe ratios in term structure models, *Working Paper, University of California, Berkeley.*

——— , 2011b, Forecasting with the term structure: The role of no-arbitrage restrictions, *Working Paper, Johns Hopkins University.*

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, p. 379–406.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2009, Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *The Journal of Finance* 64, 579–629.

Ehling, Paul, Michael Gallmeyer, Christian Heyerdahl-Larsen, and Philipp Illeditsch, 2013, Disagreement about inflation and the yield curve, Working Paper.

Fama, Eugene, 1984, The information in the term structure, *Journal of Financial Economics* 13, 509–528.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Feldhütter, Peter, 2008, Can affine models match the moments in bond yields?, working paper pp. 1–51.

——— , Christian Heyerdahl-Larsen, and Philipp Illeditsch, 2013, Expanded term structure models, Work in progress.

Garcia, Rene, 1998, Asymptotic null distribution of the likelihood ratio test in markov switching models, *International Economic Review* 39, 763–788.

Hamilton, James D., and Jing Cynthia Wu, 2012, Identification and estimation of gaussian affine term structure models, *Journal of Econometrics* 168, 315–331.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward Exchange Rates as Optimal predictors of future Spot Rates: An Econometric Analysis, *Journal of Political Economy* 88, 829–853.

Jacobs, Kris, and Lotfi Karoui, 2009, Conditional volatility in affine term-structure models: Evidence from Treasury and swap markets, *Journal of Financial Economics* 91, 288–318.

Jermann, Urban, 2013, A production-based model for the therm structure, University of Pennsylvania.

Jones, Christopher S., 2003, Nonlinear mean reversion in the short-term interest rate, *Review of Financial Studies* 16, 793–843.

Joslin, Scott, Ahn Le, and Kenneth Singleton, 2013, Why Gaussian Macro-Finance Term Structure Models Are (Nearly) Unconstrained Factor-VARs, *Journal of Financial Economics, forthcoming.*

Joslin, Scott, Marcel Priebsch, and Kenneth Singleton, 2012, Risk premiums in dynamic term structure models with unspanned macro risks, Working paper.

Joslin, Scott, Kenneth Singleton, and Haoxiang Zhu, 2011, A new perspective on gaussian dynamic term structure models, *Review of Financial Studies* 24, 926–970.

Judd, Kenneth L., 1998, *Numerical Methods in Economics* (MIT Press) 1 edn.

Kim, Don H., and Kenneth J. Singleton, 2011, Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields, *Working Paper.*

Leippold, Markus, and Liuren Wu, 2003, Design and Estimation of Quadratic Term Structure Models, *European Finance Review* 7, 47–73.

Ludvigson, Sydney, and Serena Ng, 2009, Macro Factors in Bond Risk Premia, *Review of Financial Studies* 22, 5027–5067.

Pritsker, Matt, 1998, Nonparametric density estimation and tests of continuous time interest rate models, *Review of Financial Studies* 11, 449–487.

Richard, Scott, 2013, A non-linear macroeconomic term structure model, Working Paper.

Rogers, L.C.G., 1997, The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* 7, 157–164.

Sangvinatsos, Antonios, and Jessica A. Wachter, 2005, Does the failure of the expectations hypothesis matter for long-term investors?, *Journal of Finance* 60, 179–179.

Schwartz, Eduardo S., and Anders B. Trolle, 2012, The swaption cube, Working Paper.

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment* (Princeton University Press) 1 edn.

Stanton, Richard, 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* VII, 1973–2002.

Tang, Huarong, and Yihong Xia, 2007, An international examination of affine term structure models and the expectation hypothesis, *Journal of Financial and Quantitative Analysis* 42, 41–80.

Vasicek, O., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.

Yan, Hongjun, 2008, Natural selection in financial markets: Does it work?, *Management Science* 54, 1935–1950.

A General Model

Let γ denote a nonnegative constant and $M_0(t)$ a strictly positive stochastic process with dynamics given in equation (3). The stochastic discount factor is defined as

$$M(t) = M_0(t)\left(1 + \gamma e^{-\beta' X(t)}\right)^{\alpha}, \tag{28}$$

where $\beta \in \mathcal{R}^d$ and $\alpha \in \mathcal{N}$.

We show in the next theorem that the price of a bond is a weighted average of bond prices in artificial economies that belong to the class of essentially affine Gaussian term structure models.

Theorem 2. *The price of a zero-coupon bond that matures in τ years is*

$$P^{(\tau)}(t) = \sum_{n=0}^{\alpha} w_n(t) P_n^{(\tau)}(t), \tag{29}$$

where

$$P_n^{(\tau)}(t) = e^{A_n^*(\tau) - B_n^*(\tau)' X(t)}, \tag{30}$$

$$w_n(t) = \frac{\binom{\alpha}{n}\gamma^n e^{-n\beta' X(t)}}{(1 + \gamma e^{-\beta' X(t)})^{\alpha}}. \tag{31}$$

The coefficient $A_n^*(\tau)$ and the d-dimensional vector $B_n^*(\tau)$ solve the ordinary differential equations given in equation (10) and (11).

Proof. Using the binomial expansion theorem, the stochastic discount factor in Equation (28) can be expanded as

$$M(t) = \sum_{n=0}^{\alpha} M_n(t), \tag{32}$$

where

$$M_n(t) = \binom{\alpha}{n} \gamma^n e^{-n\beta' X(t)} M_0(t). \tag{33}$$

Each summand can be interpreted as a stochastic discount factor in an artificial economy.[20] The dynamics of the strictly positive stochastic process $M_n(t)$ are

$$\frac{dM_n(t)}{M_n(t)} = -r_n(t)\, dt - \Lambda_n(t)' dW(t), \tag{34}$$

where

$$\Lambda_n(t) = \Lambda_0(t) + n\Sigma'\beta \tag{35}$$

$$r_n(t) = r_0(t) + n\beta'\kappa\left(\bar{X} - X(t)\right) - \frac{n^2}{2}\beta'\Sigma\Sigma'\beta - n\beta'\Sigma\Lambda_0(t). \tag{36}$$

Plugging in for $r_0(t)$ and $\Lambda_0(t)$, it is straightforward to show that $\Lambda_n(t)$ and $r_n(t)$ are affine functions of $X(t)$ with coefficients given in Equations (12)-(15). If $M_n(t)$ is interpreted as a stochastic discount factor of an artificial economy indexed by n then we know that bond prices in this economy belong to the class of essentially (exponential) affine Gaussian term structure models and hence

$$P_n^{(\tau)}(t) = e^{A_n^*(\tau) - B_n^*(\tau)' X(t)}, \tag{37}$$

where coefficient $A_n^*(\tau)$ and the d-dimensional vector $B_n^*(\tau)$ solve the ordinary differ-

[20] Similar expansions of the stochastic discount factor appear in Yan (2008), Dumas, Kurshev, and Uppal (2009), Bhamra and Uppal (2010), and Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013).

ential equations (10) and (11). Hence, the bond price is

$$P^{(\tau)}(t) = \sum_{n=0}^{\alpha} w_n(t) P_n^{(\tau)}(t), \tag{38}$$

where $w_n(t)$ is given in equation (31).

$\square$

Proof of Theorem 1. Set $\alpha = 1$ in Theorem 2. $\square$

Applying Ito's lemma to equation (28) leads to the dynamics of the stochastic discount factor:

$$\frac{dM(t)}{M(t)} = -r(t)\, dt - \Lambda(t)' dW(t), \tag{39}$$

where

$$\begin{aligned} r(t) = r_0(t) &+ \alpha\,(1 - s(t))\,\beta'\kappa\left(\bar{X} - X(t)\right) - \alpha(1 - s(t))\beta'\Sigma\Lambda_0(t) \\ &- \frac{\alpha}{2}\,(1 - s(t))\,(\alpha\,(1 - s(t)) + s(t))\,\beta'\Sigma\Sigma'\beta. \end{aligned} \tag{40}$$

and

$$\Lambda(t) = \Lambda_0(t) + \alpha\,(1 - s(t))\,\Sigma'\beta. \tag{41}$$

Let $\omega_n^{(\tau)}(t)$ denote the contribution of each artificial exponential affine bond price to the total bond price. Specifically,

$$\omega_n^{(\tau)}(t) = \frac{P_n^{(\tau)}(t) w_n(t)}{P^{(\tau)}(t)}. \tag{42}$$

The dynamics of the bond price $P^{(\tau)}(t)$ are

$$\frac{dP^{(\tau)}(t)}{P^{(\tau)}(t)} = \left(r(t) + \Lambda(t)' v^{(\tau)}(t)\right)\, dt + v^{(\tau)}(t)'\, dW(t), \tag{43}$$

where

$$v^{(\tau)}(t) = -\Sigma' \left(\sum_{n=0}^{\alpha} \omega_n^{(\tau)}(t) B_n^*(\tau) + \beta \left(\sum_{n=0}^{\alpha} n\, \omega_n^{(\tau)}(t) - \alpha(1 - s(t)) \right) \right). \tag{44}$$

B Equilibrium Models

In this section we show that the functional form of the state price density in equation (2) and (28) naturally comes out of several equilibrium models.[21] We need to allow for state variables that follow arithmetic Brownian motions and hence we rewrite the dynamics of the state vector in equation (1) in the slightly more general form

$$dX(t) = (\theta - \kappa X(t))\ dt + \Sigma\, dW(t), \tag{45}$$

where θ is d–dimensional and κ and Σ are $d \times d$–dimensional.

In what follows the standard consumption based asset pricing model with a representative agent power utility and log-normally distributed consumption will serve as our benchmark model. Specifically, the state price density takes the following form

$$M_0(t) = e^{-\rho t} C(t)^{-R}, \tag{46}$$

where R is the coefficient of RRA and C is aggregate consumption with dynamics

$$\frac{dC(t)}{C(t)} = \mu_C dt + \sigma_C' dW(t). \tag{47}$$

[21] Chen and Joslin (2012) provide an alternative way to solve many of these equilibrium models that is based on a nonlinear transform of processes with tractable characteristic functions.

The short rate and the market price of risk are both constant and given by

$$\Lambda_0 = R\sigma_C \tag{48}$$

$$r_0 = \rho + R\mu_C - \frac{1}{2}R(R+1)\sigma_C'\sigma_C. \tag{49}$$

Table 7 summarizes the relation between the nonlinear term structure models and the equilibrium models discussed in this section.

B.A Two Trees

Cochrane, Longstaff, and Santa-Clara (2008) study an economy in which aggregate consumption is the sum of two Lucas trees. In particular they assume that the dividends of each tree follow a geometric Brownian motion

$$dD_i(t) = D_i(t)\left(\mu_i dt + \sigma_i' dW(t)\right). \tag{50}$$

Aggregate consumption is $C(t) = D_1(t) + D_2(t)$. There is a representative agent with power utility and risk aversion R. Hence, the stochastic discount factor is

$$\begin{aligned} M(t) &= e^{-\rho t}C(t)^{-R} \\ &= e^{-\rho t}\left(D_1(t)+D_2(t)\right)^{-R} \\ &= e^{-\rho t}D_1(t)^{-R}\left(1+\frac{D_2(t)}{D_1(t)}\right)^{-R} \\ &= M_0(t)\left(1+e^{log(D_2(t))-log(D_1(t))}\right)^{-R}, \end{aligned} \tag{51}$$

where $M_0(t) = e^{-\rho t}D_1^{-R}$ and $X(t) = log\left(D_1(t)/D_2(t)\right)$. Equation (51) has the same form as the SDF in equation (28) with $\alpha \notin \mathcal{N}$. Specifically, $\gamma = 1$, $\beta = 1$, and $\alpha = -R$. Note that in this case the state variable is the log-ratio of two geometric Brownian motions and thus $\kappa = 0$. The share $s(X(t))$ and hence yields are not stationary.

B.B Multiple Consumption Goods

Models with multiple consumption goods and CES consumption aggregator naturally falls within the functional form of the SDF in equation (28). Consider a setting with two consumption goods. The aggregate output of the two goods are given by

$$dD_i(t) = D_i(t)\left(\mu_i dt + \sigma_i' dW(t)\right). \tag{52}$$

Assume that the representative agent has the following utility over aggregate consumption C,

$$u(C,t) = e^{-\rho t}\frac{1}{1-R}C^{1-R}, \tag{53}$$

where

$$C(C_1, C_2) = \left(\phi^{1-b}C_1^b + (1-\phi)^{1-b}\, C_2^b\right)^{\frac{1}{b}}. \tag{54}$$

We use the aggregate consumption bundle as numeraire, and consequently the state price density is

$$\begin{aligned} M(t) &= e^{-\rho t}C(t)^{-R} \\ &= (\phi)^{\frac{bR}{1-b}}\, e^{-\rho t} D_1(t)^{-R}\left(1+\left(\frac{1-\phi}{\phi}\right)^{1-b}\left(\frac{D_2(t)}{D_1(t)}\right)^b\right)^{-\frac{R}{b}}. \end{aligned} \tag{55}$$

After normalizing equation (55) has the same form as the SDF in equation (28) with $\alpha \notin \mathcal{N}$. Specifically, $X(t) = log(D_1(t)/D_2(t))$, $\gamma = \left(\frac{1-\phi}{\phi}\right)^{1-b}$, $\beta = b$, and $\alpha = -\frac{R}{b}$. As in the case with Two Trees, the share $s(X(t))$ and hence yields are not stationary.

B.C External Habit Formation

The utility function in Campbell and Cochrane (1999) is

$$U(C,H) = e^{-\rho t}\frac{1}{1-R}(C-H)^{1-R}, \tag{56}$$

where H is the habit level. Rather than working directly with the habit level, Campbell and Cochrane (1999) define the surplus consumption ratio $s = \frac{C-H}{C}$. The stochastic discount factor is

$$\begin{aligned} M(t) &= e^{-\rho t} C(t)^{-R} s(t)^{-R} && (57)\\ &= M_0(t) s(t)^{-R}. && (58) \end{aligned}$$

Define the state variable

$$dX(t) = \kappa\left(\bar{X} - X(t)\right)dt + b dW(t), \tag{59}$$

where $\kappa > 0$,$\sigma_c > 0$ and $b > 0$. Now let $s(t) = \frac{1}{1+e^{-\beta X(t)}}$. Note that $s(t)$ is between 0 and 1. In particular, $s(t)$ follows

$$ds(t) = s(t)\left(\mu_s(t)dt + \sigma_s(t)dW(t)\right), \tag{60}$$

where

$$\begin{aligned} \mu_s(t) &= (1-s(t))\left(\beta\kappa\left(\bar{X} - X(t)\right) + \frac{1}{2}(1-2s(t))\beta^2 b^2\right) && (61)\\ \sigma_s(t) &= (1-s(t))\beta b. && (62) \end{aligned}$$

The functional form of the surplus consumption ratio differs from Campbell and Cochrane (1999). However, note that the surplus consumption ratio is locally perfectly correlated with consumption shocks, mean-reverting and bounded between 0 and 1 just as in Campbell and Cochrane (1999). The state price density can be written as

$$M(t) = M_0(t)\left(1 + e^{-\beta X(t)}\right)^R. \tag{63}$$

The above state price density has the same form as equation (28) with parameters $\gamma = 1$, $\beta = \beta$, and $\alpha = R$. Note that the state variable X in this case is mean-reverting and therefore the share $s(X(t))$ and hence yields are stationary.

B.D Heterogeneous Beliefs

Consider an economy with two agents that have different beliefs. Let both agents have power utility with the same coefficient of relative risk aversion, R. Moreover, assume that aggregate consumption follows the dynamics in equation (47). The agents do not observe the expected growth rate and agree to disagree.[22] The equilibrium can be solved by forming the central planner problem with stochastic weight λ that captures the agents' initial relative wealth and their differences in beliefs (see Cuoco and He (1994), Basak and Cuoco (1998) and Basak (2000), for example),

$$U(C, \lambda) = \max_{\{C_1 + C_2 = C\}} \left(\frac{1}{1-R} C_1^{1-R} + \lambda \frac{1}{1-R} C_2^{1-R} \right). \tag{64}$$

Solving the above problem leads to the optimal consumption of the agents

$$C_1(t) = s(t)C(t), \tag{65}$$

$$C_2(t) = (1 - s(t))\, C(t), \tag{66}$$

[22] The model can easily be generalised to a setting with disagreement about multiple stochastic processes and learning. For instance, Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013) show that in a model with disagreement about inflation, the bond prices are weighted averages of quadratic Gaussian term structure models.

where $s(t) = \frac{1}{1+\lambda(t)^{\frac{1}{R}}}$ is the consumption share of the first agent and C is the aggregate consumption. The state price density as perceived by the first agent is

$$\begin{aligned} M(t) &= e^{-\rho t} C_1(t)^{-R} \\ &= e^{-\rho t} C(t)^{-R} s(t)^{-R} \\ &= M_0(t) \left(1 + e^{\frac{1}{R} log(\lambda(t))}\right)^{R}. \end{aligned} \tag{67}$$

This has the same form as equation (28) with $X(t) = log(\lambda(t))$, $\gamma = 1$, $\beta = -\frac{1}{R}$, and $\alpha = R$. The dynamics of the state variable is driven by the log-likelihood ratio of the two agents and consequently the share $s(X(t))$ and hence yields are not stationary.

B.E HARA Utility

Consider a pure exchange economy with a representative agent with utility $u(t,c) = \frac{e^{-\rho t}}{1-R} (C+b)^{1-R}$, where $R > 0$ and $b > 0$. We can write the SDF as

$$\begin{aligned} M(t) &= e^{-\rho t} C(t)^{-R} \\ &= e^{-\rho t} \left(C(t) + b\right)^{-R} \\ &= e^{-\rho t} C(t)^{-R} \left(1 + \frac{b}{C(t)}\right)^{-R} \\ &= M_0(t) \left(1 + e^{log(b) - log(C(t))}\right)^{-R} \end{aligned} \tag{68}$$

After normalizing equation (68) has the same form as the SDF in equation (28) with $\alpha \notin \mathcal{N}$. Specifically, $X(t) = log(b/C(t))$, $\gamma = 1$, $\beta = 1$, and $\alpha = -R$. Similarly to the model with Two Trees and multiple consumption goods, the share $s(X(t))$ and hence yields are nonstationary as the ratio $b/C(t)$ will eventually converge to zero or infinity depending on the expected growth in the economy.

C Gauss-Hermite Quadrature

While bond prices and bond yields are given in closed form conditional moments of yields and bond returns are not. However, it is straightforward to calculate conditional expectations using Gauss-Hermite polynomials because the state vector $X(t)$ is Gaussian.[23]

In this section we illustrate how to calculate the expectation of a function of Gaussian state variables. Let μ_X and Σ_X denote the conditional mean and variance of $X(s)$ at time t. Let $f(X(t))$ be a function of the state vector at time t. For instance if you want to calculate at time t the n-th uncentered moment of the bond yield with maturity τ at time u, then $f(X(u)) = \left(y^{(\tau)}(X(u))\right)^n$. Hence, the conditional expectation of $y^{(\tau)}(X(u))$ at time t is

$$E_t\left[f\left(X(u)\right)\right] = \int_{\mathcal{R}^d} f(x)\frac{1}{\left(\left(2\pi\right)^d\left|\Sigma_X\right|\right)^{0.5}}e^{-\frac{1}{2}(x-\mu_X)'\Sigma_X^{-1}(x-\mu_X)}dx. \tag{69}$$

Define $y = \sqrt{2}\sigma_X^{-1}\left(x - \mu_X\right)$ where σ_X is determined by the Cholesky decomposition $\Sigma_X = \sigma_X\sigma_X'$. Hence, we can write Equation (69) as

$$\pi^{-\frac{d}{2}}\int_{\mathcal{R}^d} f(\sqrt{2}\sigma_X y + \mu_X)e^{-y'y}dy. \tag{70}$$

Let $g(y) = f(\sqrt{2}\sigma_X y + \mu_X)$. We set $d = 3$ in the empirical section of the paper and thus the integral in Equation (70) can be approximated by the n point Gauss-Hermite quadrature

$$\int_{\mathcal{R}^d} f(\sqrt{2}\sigma_X y + \mu_X)e^{-y'y}dy \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} w_i w_j w_k g(y_1(i), y_2(j), y_3(k)), \tag{71}$$

where w_i are the weighs and $y_l(i)$ are the nodes for the n point Gauss-Hermite quadra-

[23]For more details see Judd (1998).

ture for $i = 1, .., n$ and $l = 1, .., 3$. We use $n = 4$ in equation (71).

	Three-factor Gaussian model			Three-factor nonlinear model		
κ	0.645 (0.341)	0	0	0.772 (0.326)	0	0
	0.305 (0.207)	0.0376 (0.0456)	0	0.466 (0.369)	0.0272 (0.0366)	0
	0.448 (0.197)	0.344 (0.149)	0.704 (0.296)	1.02 (0.305)	0.253 (0.117)	0.362 (0.246)
ρ_0		0.00577 (0.0873)			-0.00139 (0.0522)	
ρ_X	0.0019 (0.00873)	0.0163 (0.00247)	0.0114 (0.00448)	0.0038 (0.00176)	0.00987 (0.00111)	0.0036 (0.00364)
λ_0	0.86 (0.757)	0.159 (0.515)	0.177 (0.856)	0.356 (0.166)	-0.309 (0.307)	0.234 (0.874)
λ_X	-0.333 (0.277)	-0.0886 (0.064)	0.00254 (0.236)	-0.622 (0.222)	0.00548 (0.016)	-0.0235 (0.0447)
	-0.000635 (0.275)	0.098 (0.109)	0.396 (0.102)	0.000636 (0.329)	0.155 (0.0883)	0.259 (0.0617)
	0.34 (0.201)	-0.136 (0.1)	-0.106 (0.234)	0.512 (0.271)	-0.142 (0.113)	-0.0621 (0.122)
γ		0			0.000295 (0.00123)	
β	0	0	0	-0.866 (0.0684)	-0.449 (0.129)	0.207 (0.24)
σ		$4.63e-007$ ($9.51e-009$)			$4.57e-007$ ($8.24e-009$)	
logL		19578.6			19836.9	

Table 1: *Parameter estimates.* A standard three-factor Gaussian and a nonlinear three-factor Gaussian term structure model are estimated using the Kalman and unscented Kalman filter, respectively. This table contains the filter estimates and asymptotic standard errors (in parenthesis). The models are fitted to monthly Fama-Bliss data of one through five-year zero coupon bond yields from 1952:6 to 2011:12.

Panel A: Campbell-Shiller regression coefficients

Bond maturity	2-year	3-year	4-year	5-year
Data	-0.509	-0.849	-1.26	-1.3
	(0.532)	(0.62)	(0.659)	(0.704)
Gaussian model	-0.104	-0.23	-0.408	-0.609
Nonlinear model	-0.259	-0.53	-0.745	-0.917

Panel B: Regressing realized excess returns on expected excess retuns

	Nonlinear model				Gaussian model			
	$\alpha \times 10^3$	β	R^2	FVE	$\alpha \times 10^3$	β	R^2	FVE
One-year holding horizon								
2-year bond	-2.81	1.01	0.14	0.12	-6.64	1.77	0.16	0.12
	(3.38)	(0.36)			(3.96)	(0.55)		
3-year bond	-5.00	1.00	0.12	0.10	-11.41	1.66	0.14	0.10
	(6.41)	(0.39)			(7.33)	(0.54)		
4-year bond	-8.63	1.10	0.13	0.11	-15.72	1.66	0.15	0.12
	(8.63)	(0.39)			(9.34)	(0.51)		
5-year bond	-10.66	1.09	0.13	0.11	-16.54	1.52	0.15	0.12
	(9.97)	(0.39)			(10.28)	(0.46)		
Two-year holding horizon								
3-year bond	-5.42	1.04	0.17	0.14	-8.85	1.42	0.12	0.09
	(7.15)	(0.47)			(9.18)	(0.72)		
4-year bond	-11.60	1.10	0.18	0.14	-15.26	1.36	0.11	0.09
	(12.86)	(0.48)			(15.65)	(0.68)		
5-year bond	-18.05	1.15	0.20	0.16	-20.32	1.32	0.13	0.09
	(16.38)	(0.46)			(18.88)	(0.60)		
Three-year holding horizon								
4-year bond	-6.44	0.98	0.17	0.13	-8.91	1.18	0.10	0.07
	(10.42)	(0.56)			(14.02)	(0.85)		
5-year bond	-11.37	0.95	0.16	0.11	-13.16	1.05	0.09	0.05
	(18.26)	(0.56)			(23.17)	(0.78)		
Four-year holding horizon								
5-year bond	-8.66	1.00	0.19	0.14	-10.35	1.09	0.10	0.05
	(12.76)	(0.63)			(17.55)	(0.93)		

Table 2: *Excess return regressions.* Panel A shows the coefficients ϕ^τ from the regressions $y(t+1,\tau-1) - y(t,\tau) = const + \phi^\tau[\frac{y(t,\tau)-y(t,1)}{\tau-1}]$ + residual, where $y(t,\tau)$ is the zero-coupon yield at time t of a bond maturing at time $t+\tau$ (τ and t are measured in years). The actual coefficients are calculated using monthly Fama-Bliss data of one through five-year zero coupon bond yields from 1952:6 to 2011:12. For each model the coefficient is based on one simulated sample path of 1,000,000 months. Panel B shows regression coefficients from a regression of realized excess returns on expected excess returns in sample. FVA is $1 - \frac{\frac{1}{T}\sum_{t=1}^{T}(RX^\tau_{t,t+n} - E_t(RX^\tau_{t,t+n}))^2}{\frac{1}{T}\sum_{t=1}^{T}(RX^\tau_{t,t+n} - \overline{RX^\tau})^2}$, where $RX^\tau_{t,t+n}$ is the n-year excess return on a bond with maturity τ. For both panels standard errors in parentheses are Hansen and Hodrick (1980) with number of lags equal to the number of overlapping months. In Panel A standard errors for both models are small and thus omitted.

Panel A: EGARCH yield vol. regressed on model-implied yield vol.

	$\alpha \times 10^3$	β	corr.	$\overline{\alpha} \times 10^3$	$\overline{\beta}$	FVE
1-year bond	0.54	0.87	72.7	0.88	0.71	50.4
	(0.44)	(0.10)				
2-year bond	0.91	0.77	74.6	0.88	0.68	52.1
	(0.35)	(0.08)				
3-year bond	0.92	0.76	74.5	0.91	0.67	51.2
	(0.34)	(0.09)				
4-year bond	1.01	0.77	70.7	0.91	0.65	41.5
	(0.34)	(0.09)				
5-year bond	1.10	0.71	74.7	0.88	0.65	46.8
	(0.29)	(0.08)				

Panel B: Distribution of one-month ahead conditional yield volatility (bps/month)

	mean	0.01	0.05	0.25	median	0.75	0.95	0.99
GARCH(1,1)								
1-year bond	39	13	15	21	31	45	96	164
2-year bond	37	16	18	24	33	42	80	130
3-year bond	36	14	16	25	33	41	72	111
4-year bond	36	14	15	24	33	42	72	111
5-year bond	33	13	16	24	30	38	65	89
Gaussian								
1-year bond	46	46	46	46	46	46	46	46
2-year bond	41	41	41	41	41	41	41	41
3-year bond	38	38	38	38	38	38	38	38
4-year bond	36	36	36	36	36	36	36	36
5-year bond	34	34	34	34	34	34	34	34
Nonlinear								
1-year bond	39	26	26	27	28	36	100	138
2-year bond	36	25	25	26	28	36	88	117
3-year bond	35	24	24	25	27	36	80	106
4-year bond	33	23	23	24	26	35	74	98
5-year bond	31	21	22	23	25	32	70	93

Table 3: *Volatility regressions.* Panel A shows regression coefficients α and β from regressing an EGARCH(1,1) estimate of monthly conditional volatility on model-implied conditional one-month ahead volatility in sample. Panel A also shows the regression coefficients $\overline{\alpha}$ and $\overline{\beta}$ resulting from simulating 100,000 months from the nonlinear model, estimating an EGARCH(1,1) volatility from the simulated yields, and regressing the estimated EGARCH volatility on model-implied volatility. FVE is $1 - \frac{\frac{1}{T}\sum_{t=1}^{T}(Var_t^{EGARCH}(y^\tau(t+1)-y^\tau(t))-[\overline{\alpha}+\overline{\beta}Var_t^{nonlinear}(y^\tau(t+1)-y^\tau(t))])^2}{\frac{1}{T}\sum_{t=1}^{T}(Var_t^{EGARCH}(y^\tau(t+1)-y^\tau(t))-\overline{Var_t^{EGARCH}(y^\tau(t+1)-y^\tau(t))})^2}$, where $Var_t^i(y^\tau(t+1)-y^\tau(t))$ is conditional volatility at time t for a bond with maturity τ in model i. Standard errors in parentheses are Hansen and Hodrick (1980) with 12 lags. Panel B shows mean and quantiles of the distribution of conditional volatility in basis points per month.

	Variance of conditional expectation		
Bond maturity	Expectation	Expectation projected on yields	Ratio
Panel A: Instantaneous excess returns			
Gaussian three-factor model			
2	0.99	0.99	1
3	1.95	1.95	1
4	3.27	3.27	1
5	5.01	5.01	1
Nonlinear three-factor model			
2	2.14	1.70	0.79
3	3.74	2.88	0.77
4	5.40	4.02	0.74
5	7.04	5.04	0.71
Panel B: Yearly excess returns			
Gaussian three-factor model			
2	0.14	0.14	1
3	0.47	0.47	1
4	1.02	1.02	1
5	1.83	1.83	1
Nonlinear three-factor model			
2	0.36	0.31	0.86
3	1.03	0.86	0.84
4	1.80	1.48	0.82
5	2.68	2.14	0.80

Table 4: *Model-implied population properties of excess returns.* This table reports population properties of instantaneous and annual log excess returns of a n-year bond. Excess returns are calculated by subtracting the short rate and the one-year return on a one-year bond, respectively. Conditional expectations are calculated in the model and compared to the conditional expectations derived by linearly projecting the model-implied expectations onto the five model-implied yields. In the Gaussian model the former and the latter are the same. Variances are in percent squared.

Inflation	Industrial production growth	Duffee(2011) hidden factor	R^2
-0.04 (-2.26)			0.06
	-0.00 (-0.07)		0.00
		0.34 (0.55)	0.00
-0.04 (-2.26)	0.00 (0.10)		0.06
-0.03 (-1.63)	-0.00 (-0.12)	0.47 (0.79)	0.05

Table 5: *Projections of the linearly hidden risk premium component in the three-factor nonlinear model.* There is a nonlinear relation between expected excess returns and yields in the nonlinear model. We regress model-implied one-year expected excess returns averaged across the 2-, 3-, 4-, and 5-year bond on model-implied yields. The residual from the regression is the linearly hidden part of the average expected excess returns. We then regress the hidden part on CPI inflation and industrial production growth over the next 12 months (log changes over the next 12 months). The Duffee (2011) hidden factor is calculated by downloading the smoothed risk premium factor from Greg Duffee's webpage and taking the residual from the projection onto yields with maturities ranging from one to five years. The data sample when the hidden factor is included is 1964:1-2007:12. The t-statistics in parentheses are based on Hansen and Hodrick (1980) standard errors with 12 lags.

Panel A: Principal components of yields

	1st PC	1st & 2nd PC	1st, 2nd, & 3rd PC
Data	0.9906	0.9995	0.9998
Gaussian model	0.9967	0.9999	1.0000
Nonlinear model	0.9959	0.9998	1.0000

Panel B: Regressing yields on principal components

	Maturity	1st PC	2nd PC	3rd PC	4th PC	5th PC
Data	1	0.46	-0.73	-0.46	0.19	0.06
	2	0.46	-0.22	0.50	-0.61	-0.35
	3	0.45	0.12	0.48	0.29	0.68
	4	0.44	0.36	0.02	0.56	-0.60
	5	0.43	0.53	-0.55	-0.44	0.21
Gaussian model	1	0.45	-0.74	-0.45	0	0
	2	0.45	-0.22	0.48	0	0
	3	0.45	0.12	0.48	0	0
	4	0.44	0.35	0.04	0	0
	5	0.44	0.52	-0.58	0	0
Nonlinear model	1	0.46	-0.70	-0.48	0.24	0.08
	2	0.46	-0.25	0.42	-0.63	-0.40
	3	0.45	0.09	0.51	0.18	0.71
	4	0.44	0.36	0.11	0.60	-0.56
	5	0.43	0.56	-0.57	-0.39	0.16

Table 6: *Cross-sectional fit of three-factor models.* Panel A shows the contribution of the first three principal components to the total variation in bond yields. Principal components are constructed from a panel of constant-maturity zero-coupon bond yields with maturities ranging from one to five years. Panel B shows the slope coefficients from the regressions of each yield on all five principal components and a constant. The actual coefficients are computed using monthly Fama-Bliss data of one through five-year zero coupon bond yields from 1952:6 to 2011:12. For each model the coefficient is based on one simulated sample path of 1,000,000 months.

Model	N	d	X	α	γ	β	Stationary
Two trees	1	2	$log\,(D_1(t)/D_2(t))$	$-R$	1	1	No
Multiple consumption goods	1	2	$log\,(D_1(t)/D_2(t))$	$-\frac{R}{b}$	$\left(\frac{1-\phi}{\phi}\right)^{1-b}$	b	No
External habit formation	1	1	X	R	1	β	Yes
Heterogeneous beliefs	1	1	$log\,(\lambda(t))$	R	1	$-\frac{1}{R}$	No
HARA utility	1	1	$log\,(b/C(t))$	$-R$	1	1	No

Table 7: *Equilibrium models.* The table shows various equilibrium models and how they map into the nonlinear term structure models.

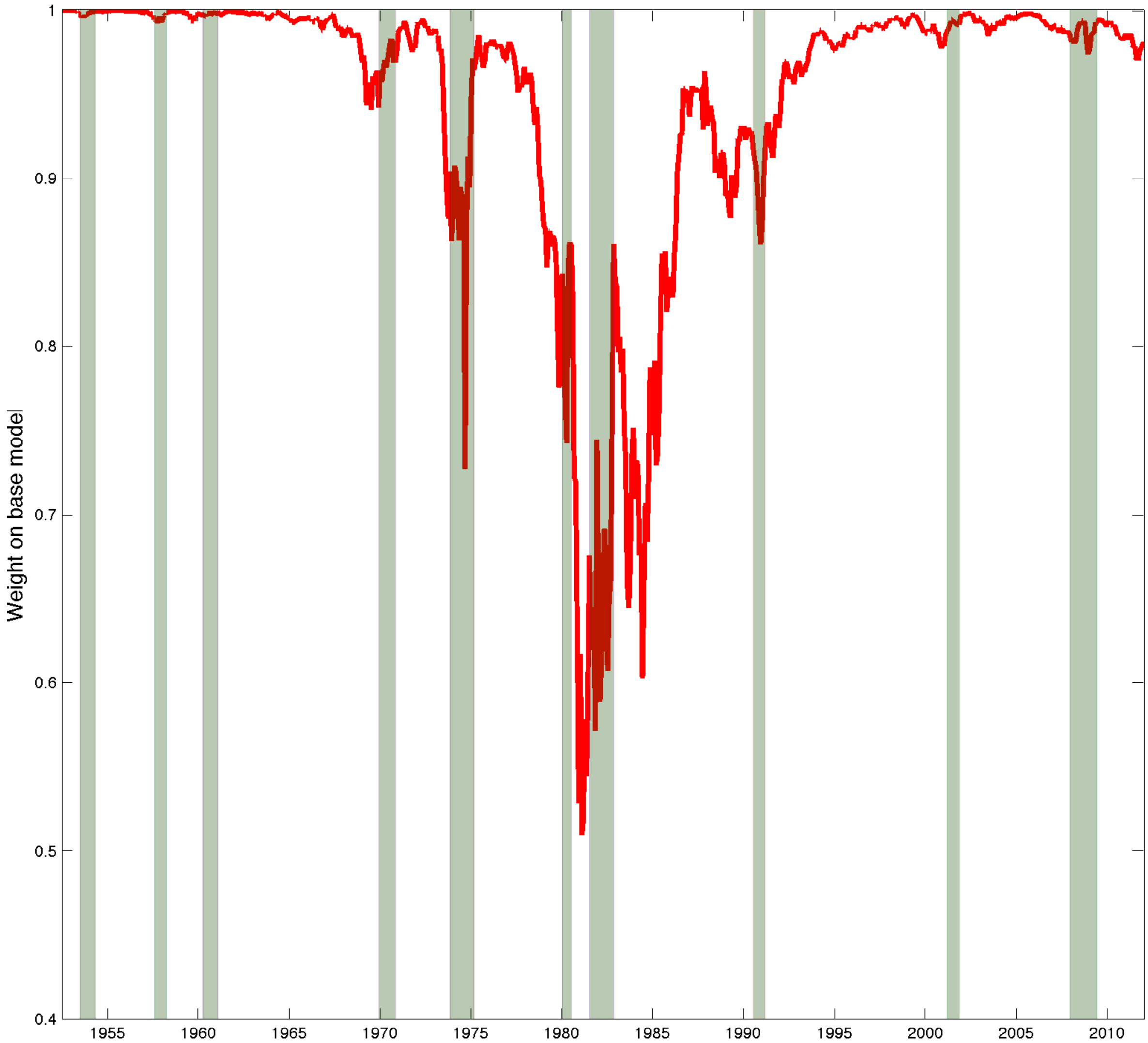

Figure 1: *The weight on the base model.* The bond price in the nonlinear model is $P^{(\tau)}(t) = s(t)P_0^{(\tau)}(t)+(1-s(t))P_1^{(\tau)}(t)$, where $P_0^{(\tau)}(t)$ is the bond price in the standard Gaussian model. $s(t)$ is a weight between 0 and 1 and the figure plots this weight. The shaded areas show NBER recessions.

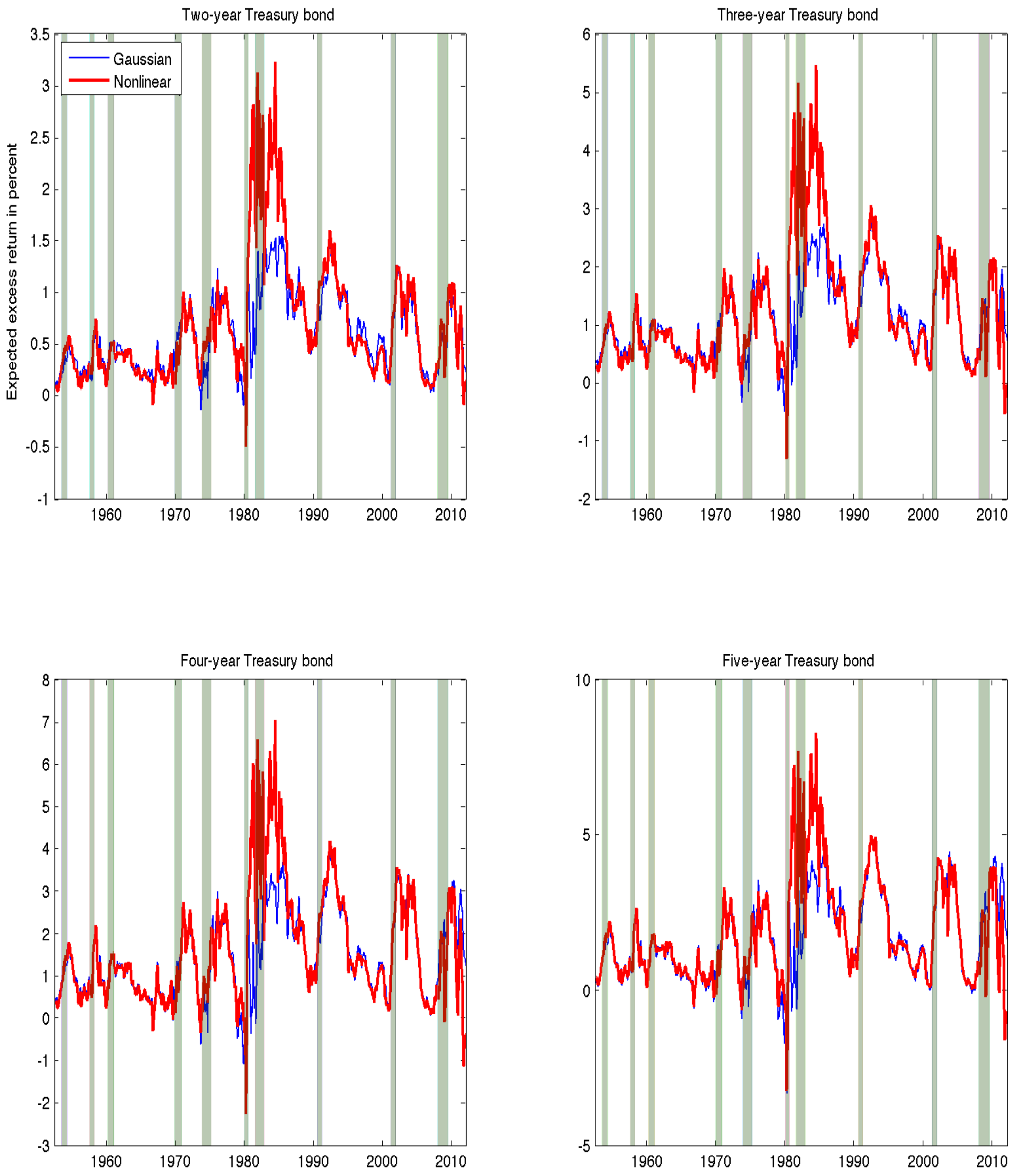

Figure 2: *Expected excess returns.* The graphs show the expected one year log excess returns of zero-coupon Treasury bonds with maturities of 2, 3, 4, and 5 years. The thin blue lines show expected excess returns in the three-factor Gaussian model and the thick red lines show expected excess returns in the three-factor nonlinear model. The shaded areas show NBER recessions.

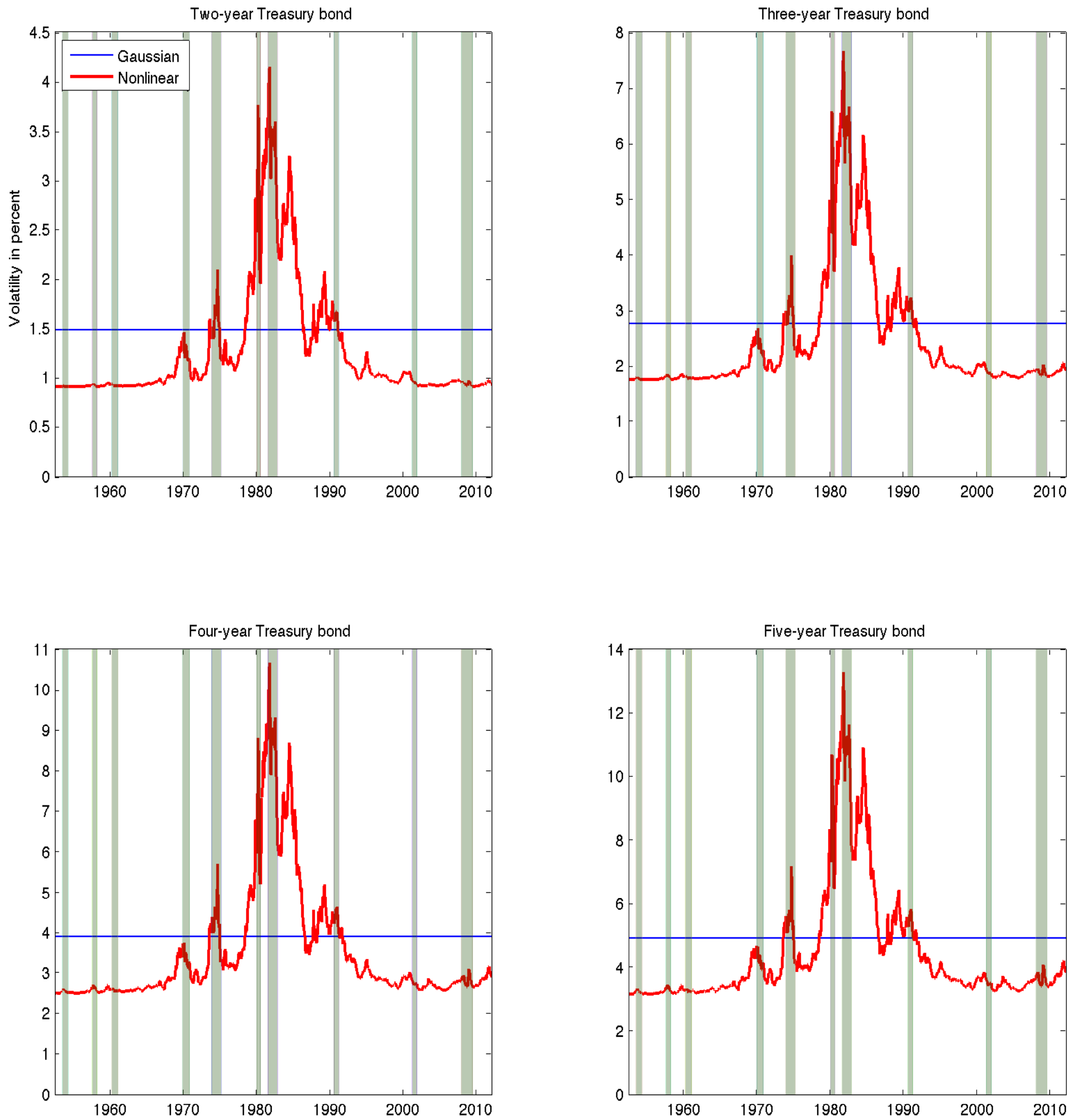

Figure 3: *Volatility of excess returns.* The graphs show the conditional volatility of one year log excess returns for zero-coupon Treasury bonds with maturities of 2, 3, 4, and 5 years. The thin blue lines show volatilities in the three-factor Gaussian model and the thick red lines show volatilities in the three-factor nonlinear model. The shaded areas show NBER recessions.

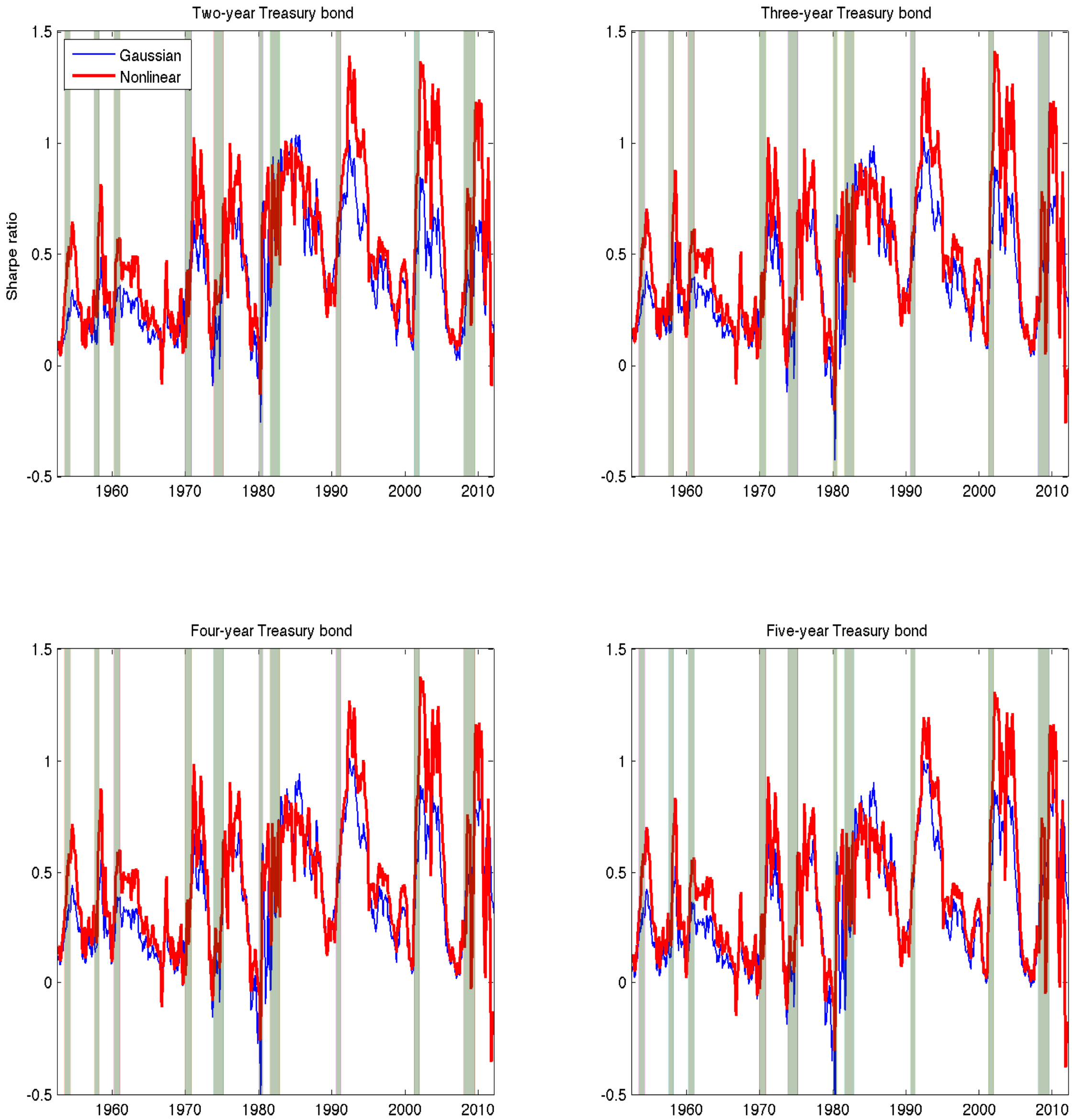

Figure 4: *Sharpe ratio.* The conditional Sharpe ratio for a bond is defined as $E_t\left[rx_{t,t+12}^{(\tau)}\right] / \sqrt{Var_t\left[rx_{t,t+12}^{(\tau)}\right]}$ where $rx_{t,t+12}^{(\tau)}$ is the one-year log excess return on a bond with maturity τ. The graphs shows the conditional Sharpe ratio for the 2-, 3-, 4- and 5-year bonds. The shaded areas show NBER recessions.