

Estimating Shadow-Rate Term Structure Models with Near-Zero Yields

Jens H. E. Christensen[†]

and

Glenn D. Rudebusch

Federal Reserve Bank of San Francisco

101 Market Street, Mailstop 1130

San Francisco, CA 94105

Abstract

Standard Gaussian term structure models have often been criticized for not ruling out negative nominal interest rates, but this flaw has been especially conspicuous with interest rates near zero in many countries. We provide a tractable means to estimate an alternative Gaussian shadow-rate dynamic term structure model that enforces the zero lower bound on bond yields. We illustrate this model by estimating one-, two-, and three-factor shadow-rate models on a sample of positive and near-zero Japanese bond yields. We find that the level of the shadow rate is sensitive to model fit and specification, including the number of factors employed.

JEL Classification: G12, E43, E52, E58.

Keywords: Term structure modeling, zero lower bound, monetary policy.

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System. We thank Leo Krippner for helpful comments.

[†]Corresponding author. E-mail: jens.christensen@sf.frb.org

This version: March 6, 2013.

1 Introduction

Nominal yields on government debt in several countries have fallen very near their zero lower bound (ZLB). Notably, yields on Japanese government bonds of various maturities have been near zero since 1996. Similarly, many U.S. Treasury rates edged down quite close to zero in the years following the financial crisis in late 2008. Accordingly, understanding how to model the term structure of interest rates when some of those interest rates are near the ZLB is an issue that commands attention both for bond portfolio pricing and risk management and for macroeconomic and monetary policy analysis. Unfortunately, the workhorse representation in finance for bond pricing—the affine Gaussian dynamic term structure model—ignores the ZLB and routinely places positive probabilities on the occurrence of future negative interest rates. This counterfactual feature results from ignoring the existence of a readily available currency for transactions. In the real world, an investor always has the option of holding cash, and the zero nominal yield of cash will dominate any security with a negative yield.¹

To recognize the option value of currency in bond pricing, Black (1995) introduced the notion of a “shadow short rate,” which is driven by fundamentals and can be positive or negative. The observed short rate equals the shadow short rate except that the former is bounded below by zero. While Black’s (1995) use of a shadow short rate to account for the presence of currency holds much intuitive appeal, it has rarely been used. In part, this infrequency reflects the fact that interest rates in many countries have long been some distance above zero, so the Gaussian positive probabilities on negative interest rates are negligible and unlikely to be an important determinant in bond pricing. In recent years, with yields around the world at historic lows, this rationale has evaporated. However, a second factor limiting the adoption of the shadow-rate structure has been the difficulty in estimating these models using the requisite computationally intensive numerical methods. For example, among the handful of shadow-rate studies, Kim and Singleton (2012), henceforth KS, estimate a shadow-rate dynamic term structure model using a sample of near-zero Japanese government bond yields. Unfortunately, the KS analysis is limited to a model with only two pricing factors because the numerical methods required for shadow-rate models with more than two factors are computationally prohibitive. This practical shortcoming is potentially quite serious given the prevalence of higher-dimensional bond pricing models in research and industry.²

To overcome the practical computational constraints imposed by shadow-rate modeling,

¹Actually, the ZLB can be a somewhat soft floor. The nonnegligible costs of transacting in and holding large amounts of currency have allowed government bond yields to push slightly below zero in a few countries, notably in Denmark recently. To account for institutional currency frictions in our analysis, we could replace the zero lower bound on yields with some appropriate, possibly time-varying, negative epsilon.

²Indeed, KS suggest that the shadow-rate model results of Ueno, Baba, and Sakurai (2006) are influenced by their use of a *one*-factor shadow-rate model that may not be flexible enough to fit their sample of Japanese data. Similarly, the KS results with two-factor models may not generalize to higher-order models.

Krippner (2012) presents a shadow-rate framework that allows for tractable estimation of dynamic term structure shadow-rate models with more than two factors. In this paper, we implement this new framework. Specifically, we combine the Krippner (2012) shadow-rate framework with the Gaussian arbitrage-free Nelson-Siegel (AFNS) models introduced in Christensen, Diebold, and Rudebusch (2011), henceforth CDR. The AFNS model class provides a flexible, robust structure for dynamic term structure modeling that has performed well on a variety of yield samples by providing good fit with tractable estimation. Furthermore, as we show in this paper, even when combined with the Krippner framework, the AFNS specification of the pricing factor dynamics leads to analytical formulas for the instantaneous shadow forward rates. This new analytical result demonstrates that empirical implementation of higher-order shadow-rate models can be straightforward.

To demonstrate the application of the new shadow-rate AFNS model and assess its empirical performance, we apply it to Japanese term structure data, which are of special interest because they include a long period of near-zero yields. In particular, we estimate one-, two-, and three-factor versions of our shadow-rate AFNS model on two samples of Japanese term structure data. First, we estimate the models on the exact same data set analyzed in KS, which is a weekly sample from January 1995 through March 2008. We use this sample to answer two questions: Does the two-factor shadow-rate AFNS model provide empirical results similar to the KS two-factor shadow-rate model when both are estimated on the same data set? We find in terms of in-sample fit and the shadow-rate process that the two models look quite close. Next, are the KS two-factor model results robust to the number of factors used in estimation? With our tractable shadow-rate estimation, we can easily examine whether the use of one, two, or three factors affects the shadow rate results. In this regard, we find that, during periods when the Japanese shadow short rate is mainly negative (the years from 2001 to 2005), there is notable disagreement about the value of shadow rates across models with different numbers of factors. This sensitivity to model specification suggests that care must be taken when using the level of the negative shadow rate as a measure of the stance of monetary policy when yields are near the zero boundary, as noted by Krippner (2013). Importantly, though, the shadow-rate paths are highly correlated across models. Thus, in terms of qualitative statements about the general *direction* of the shadow rate from one period to another, the models would tend to agree and such directional inference could be made with some robustness.

Unfortunately, the KS data sample ends in early 2008, before the recent global financial crisis, when interest rates across a wide range of maturities in Japan and in other countries fell to near zero. To provide up-to-date results on this important period, we also estimate the shadow-rate models on a second, longer sample of Japanese yields that extends the KS sample through November 2012. Here again, we see some variation across the estimated one-,

two-, and three-factor shadow-rate AFNS models; however, in all cases, the shadow rates are not as negative from 2009 to 2012 as during the 2001-2005 episode.

The rest of the paper is structured as follows. Section 2 introduces the shadow-rate framework used in KS and Krippner (2012), while Section 3 details our shadow-rate AFNS model. Section 4 describes the Japanese yield data used by KS and our extension thereof. Section 5 presents our empirical findings for one-, two-, and three-factor shadow-rate models for the KS data, and Section 6 contains our findings for the extended sample. Finally, Section 7 concludes. Two appendices provide technical details on option pricing and model estimation.

2 The Shadow-Rate Modeling Framework

In this section, we briefly summarize shadow-rate Gaussian term structure models that take into account the zero lower bound on yields. We first describe the approach of Black (1995) and its implementation by KS. We then consider the shadow-rate framework introduced in Krippner (2012).

2.1 The Black Shadow-Rate Model

The concept of a shadow interest rate as a modeling tool to account for the ZLB can be attributed to Black (1995). He noted that the observed nominal short rate will be nonnegative because currency is a readily available asset to investors that carries a nominal interest rate of zero. Therefore, the existence of currency sets a zero lower bound on yields.

To account for this ZLB, Black postulated as a modeling tool using a shadow short rate, s_t , that is unconstrained by the ZLB. The usual observed instantaneous risk-free rate, r_t , which is used for discounting cash flows when valuing securities, is then given by the greater of the shadow rate or zero:

$$r_t = \max\{0, s_t\}.$$

Accordingly, as s_t falls below zero, the observed r_t simply remains at the zero bound.

While Black (1995) described circumstances under which the zero bound on nominal yields might be relevant, he did not provide specifics for implementation. Gorovoi and Linetsky (2004) derive bond price formulas for the case of one-factor Gaussian and square-root shadow-rate models.³ Unfortunately, their results do not extend to multidimensional models. Instead, the small set of previous research on shadow-rate models has relied on numerical methods for pricing.⁴ However, in light of the computational burden of these methods, there have been

³Ueno et al. (2006) use these formulas when calibrating a one-factor Gaussian model to a sample of Japanese government bond yields.

⁴Both KS and Bomfim (2003) use finite-difference methods to calculate bond prices, while Ichiue and Ueno (2007) employ interest rate lattices.

only two previous estimations of multifactor shadow-rate models: Ichiue and Ueno (2007) and KS. Both of these studies undertake a full maximum likelihood estimation of their two-factor Gaussian shadow-rate models on Japanese bond yield data using the extended Kalman filter and numerical optimization.

Relative to Ichiue and Ueno (2007), KS use more flexible models and a longer sample; therefore, we use KS as a benchmark for comparison with our results. Here we provide details of their model. Specifically, the KS two-factor shadow-rate affine Gaussian model, denoted as the B-AG2 model (for Black, affine Gaussian), has a shadow-rate process defined as

$$s_t(X_t) = \rho + X_t^1 + X_t^2.$$

The dynamics of the state variables $X_t = (X_t^1, X_t^2)$ under the objective probability P -measure are assumed to have the following structure:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 \\ \kappa_{21}^P & \kappa_{22}^P \end{pmatrix} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \end{pmatrix}.$$

The market prices of risk are flexible extended affine, as per Cheridito, Filipović, and Kimmel (2007),

$$\lambda(X_t) = \lambda_a + \Lambda_b X_t.$$

As a consequence, the dynamics of the state variables used for pricing under the risk-neutral Q -measure have the following structure:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^Q & \kappa_{12}^Q \\ \kappa_{21}^Q & \kappa_{22}^Q \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dX_t^{2,Q} \end{pmatrix}.$$

KS derive the partial differential equation (PDE) that bond prices must satisfy under the restriction that the risk-free rate used for discounting is the greater of the shadow rate or zero,

$$\frac{\partial P}{\partial \tau} - \frac{1}{2} \text{tr} \left(\frac{\partial^2 P}{\partial x \partial x} \Sigma \Sigma' \right) - \frac{\partial P}{\partial x} K^Q (\theta^Q - x) + \max\{0, s(x)\} P = 0, \quad P(0, x) = 1.$$

KS solve this PDE using a finite-difference method. Unfortunately, for more than two factors, such numerical methods render it impractical if not impossible to solve the associated higher-dimensional PDE systems within a reasonable time. This is a severe limitation to estimating shadow-rate models since the bond pricing literature has focused on models with at least three factors driving bond yields.

2.2 The Krippner Shadow-Rate Framework

To overcome the curse of dimensionality that limits implementation of the Black shadow-rate framework, Krippner (2012) suggested an alternative shadow-rate approach that, in the case of Gaussian factor dynamics, holds the possibility of being as easy to estimate as the corresponding non-shadow-rate model. In particular, shadow-rate modeling with more than two state variables would seem to be tractable. In the rest of this subsection, we outline Krippner’s general framework.⁵

First, assume the existence of shadow zero-coupon bonds that are priced such that their expected return under the risk-neutral pricing measure always equals s_t , which is the instantaneous risk-free shadow rate that can go negative. Let the price of such a bond at time t that matures at time T be denoted by $P(t, T)$. Furthermore, assume there exists an alternative set of zero-coupon bond prices, $\underline{P}(t, T)$, with nonnegative yields that will actually be observed.⁶ Finally, assume there is currency in circulation with no transaction costs and a constant nominal value.

Consider the shadow bond at time t with the shortest (say, overnight) maturity available, δ . Investors can either choose to buy the zero-coupon bond at price $P(t, t + \delta)$ and receive one unit of currency the following day or just hold the currency. As a consequence, the price of the observed overnight claim must be capped at 1:

$$\begin{aligned}\underline{P}(t, t + \delta) &= \min\{1, P(t, t + \delta)\} \\ &= P(t, t + \delta) - \max\{P(t, t + \delta) - 1, 0\}.\end{aligned}$$

That is, the availability of currency implies that the overnight claim has a value equal to the zero-coupon shadow bond price minus the value of a call option on the zero-coupon shadow bond with a strike price of 1. Based on this, Krippner argues that a regular zero-coupon bond with maturity at $T + \delta$ is equivalent to the corresponding shadow bond minus the value of a European call option written on the shadow bond with maturity at T and a strike price of 1. As a consequence, he introduces the following auxiliary bond price equation

$$P_a(t, T + \delta) = P(t, T + \delta) - C(t, T, T + \delta; 1), \tag{1}$$

where $C(t, T, T + \delta; 1)$ is the value of a European call option at time t with maturity T and

⁵As described in Krippner (2012), there is a small, subtle difference between his framework and the Black framework. The Black framework enforces the nonnegativity of the short rate at each point in time, while the Krippner framework enforces a nonnegative forward rate curve at each point in time.

⁶The modeling approach with unobserved, or “shadow,” components has an analogy in the corporate credit literature. There, it is frequently assumed that the asset value process of a firm exists but is unobserved. Instead, prices of the firm’s equity and corporate debt, which can be interpreted as derivatives written on the firm’s assets (see Merton 1974), are used to draw inferences about the asset value process.

strike price 1 written on the shadow discount bond maturing at $T + \delta$. Importantly, it should be noted that $P_a(t, T + \delta)$ is not identical to the bond price $\underline{P}(t, T)$ whose yield observes the zero lower bound.

The key insight of Krippner (2012) is that the last incremental forward rate of any bond will be nonnegative due to the future availability of currency in the immediate time prior to its maturity. By letting $\delta \rightarrow 0$, he takes this idea to its continuous limit, which identifies the corresponding nonnegative instantaneous forward rate:

$$\underline{f}(t, T) = \lim_{\delta \rightarrow 0} \left[-\frac{d}{d\delta} P_a(t, T + \delta) \right]. \quad (2)$$

Now, the discount bond prices whose yields observe the zero lower bound are defined as

$$\underline{P}(t, T) = e^{-\int_t^T \underline{f}(t, s) ds}. \quad (3)$$

Since the observed discount bond prices defined in this way differ from the auxiliary bond price $P_a(t, T)$ defined in equation (1) and used in the construction of the nonnegative forward rate in equation (2), the Krippner framework could be viewed as not fully internally consistent. However, as we will see in the following, the auxiliary bond price drops out of the calculations, and we are left with formulas for the nonnegative forward rate, $\underline{f}(t, T)$, that are solely determined by the properties of the shadow rate process s_t . Specifically, Krippner (2012) shows that

$$\underline{f}(t, T) = f(t, T) + z(t, T),$$

where $f(t, T)$ is the instantaneous forward rate on the shadow bond, which may go negative, while $z(t, T)$ is given by

$$z(t, T) = \lim_{\delta \rightarrow 0} \left[\frac{d}{d\delta} \left\{ \frac{C(t, T, T + \delta; 1)}{P(t, T + \delta)} \right\} \right].$$

In addition, it holds that the observed instantaneous risk-free rate is

$$r_t = \max\{0, s_t\}$$

as in the Black (1995) model.

Finally, yield-to-maturity is defined the usual way as

$$\begin{aligned}
\underline{y}(t, T) &= \frac{1}{T-t} \int_t^T \underline{f}(t, s) ds \\
&= \frac{1}{T-t} \int_t^T f(t, s) ds + \frac{1}{T-t} \int_t^T \lim_{\delta \rightarrow 0} \left[\frac{\partial}{\partial \delta} \frac{C(t, s, s + \delta; 1)}{P(t, s)} \right] ds \\
&= y(t, T) + \frac{1}{T-t} \int_t^T \lim_{\delta \rightarrow 0} \left[\frac{\partial}{\partial \delta} \frac{C(t, s, s + \delta; 1)}{P(t, s)} \right] ds.
\end{aligned}$$

It follows that bond yields constrained at the ZLB can be viewed as the sum of the yield on the unconstrained shadow bond, denoted $y(t, T)$, which is modeled using standard tools, and an add-on correction term derived from the price formula for the option written on the shadow bond that provides an upward push to deliver the higher nonnegative yields actually observed. Importantly, the result above is general and applies to any assumptions made about the dynamics of the shadow-rate process. In the next section, we consider a Gaussian model that leads to tractable formulas for bond yields in the Krippner shadow-rate framework.

3 The Shadow-Rate AFNS Model

In this section, we show how the Krippner modeling framework can be implemented. To model the risk-free shadow rate, we employ the affine arbitrage-free class of Nelson-Siegel term structure models derived in CDR. This class of models is very tractable to estimate and has good in-sample fit and out-of-sample forecast accuracy.⁷ Here, we extend the AFNS model to incorporate a nonnegativity constraint on observed yields.

3.1 The Standard AFNS(3) Model

We first briefly describe the standard three-factor AFNS(3) model, which ignores the ZLB on yields. In this class of models, the risk-free rate, which we take to be the potentially unobserved shadow rate, is given by

$$s_t = X_t^1 + X_t^2,$$

⁷See, for example, the discussion and references in Diebold and Rudebusch (2013).

while the dynamics of the state variables (X_t^1, X_t^2, X_t^3) used for pricing under the Q -measure have the following structure:⁸

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dX_t^{2,Q} \\ dX_t^{3,Q} \end{pmatrix}. \quad (4)$$

The AFNS model dynamics under the Q -measure may appear restrictive, but CDR show this structure coupled with general risk pricing provides a very flexible modeling structure. Indeed, CDR demonstrate that this specification implies zero-coupon bond yields that have the popular Nelson and Siegel (1987) factor loading structure,

$$y(t, T) = X_t^1 + \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)} \right) X_t^2 + \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)} \right) X_t^3 - \frac{A(t, T)}{T-t}.$$

The yield function also contains a yield-adjustment term, $\frac{A(t, T)}{T-t}$, that is time invariant and depends only on the maturity of the bond. CDR provide an analytical formula for this term, which under our identification scheme is entirely determined by the volatility matrix.

The corresponding instantaneous forward rates are given by

$$f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) = X_t^1 + e^{-\lambda(T-t)} X_t^2 + \lambda(T-t) e^{-\lambda(T-t)} X_t^3 + A^f(t, T), \quad (5)$$

where the yield-adjustment term in the instantaneous forward rate function is given by

$$\begin{aligned} A^f(t, T) &= -\frac{\partial A(t, T)}{\partial T} \\ &= -\frac{1}{2} \sigma_{11}^2 (T-t)^2 - \frac{1}{2} (\sigma_{21}^2 + \sigma_{22}^2) \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right)^2 \\ &\quad - \frac{1}{2} (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \left[\frac{1}{\lambda^2} - \frac{2}{\lambda^2} e^{-\lambda(T-t)} - \frac{2}{\lambda} (T-t) e^{-\lambda(T-t)} \right. \\ &\quad \left. + \frac{1}{\lambda^2} e^{-2\lambda(T-t)} + \frac{2}{\lambda} (T-t) e^{-2\lambda(T-t)} + (T-t)^2 e^{-2\lambda(T-t)} \right] \\ &\quad - \sigma_{11} \sigma_{21} (T-t) \frac{1 - e^{-\lambda(T-t)}}{\lambda} \\ &\quad - \sigma_{11} \sigma_{31} \left[\frac{1}{\lambda} (T-t) - \frac{1}{\lambda} (T-t) e^{-\lambda(T-t)} - (T-t)^2 e^{-\lambda(T-t)} \right] \\ &\quad - (\sigma_{21} \sigma_{31} + \sigma_{22} \sigma_{32}) \left[\frac{1}{\lambda^2} - \frac{2}{\lambda^2} e^{-\lambda(T-t)} - \frac{1}{\lambda} (T-t) e^{-\lambda(T-t)} + \frac{1}{\lambda^2} e^{-2\lambda(T-t)} \right. \\ &\quad \left. + \frac{1}{\lambda} (T-t) e^{-2\lambda(T-t)} \right]. \end{aligned}$$

⁸We have fixed the mean under the Q -measure at zero and assumed a lower triangular structure for the volatility matrix, which comes at no loss of generality, as described by CDR.

3.2 Bond Option Prices

To apply the Krippner framework, we need the analytical formula for the price of the call option written on the shadow bond described above.

From standard asset pricing theory it follows that the value of a European call option with maturity T and strike price K written on the zero-coupon bond maturing at $T + \delta$ is given by

$$C(t, T, T + \delta; K) = E_t^Q \left[e^{-\int_t^T s_u du} \max\{P(T, T + \delta) - K, 0\} \right].$$

Calculations provided in Appendix A show that the value of the European call option within the AFNS(3) model is given by⁹

$$C(t, T, T + \delta; K) = P(t, T + \delta)\Phi(d_1) - KP(t, T)\Phi(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{P(t, T + \delta)}{P(t, T)K}\right) + \frac{1}{2}v(t, T, T + \delta)}{\sqrt{v(t, T, T + \delta)}} \quad \text{and} \quad d_2 = d_1 - \sqrt{v(t, T, T + \delta)}$$

with

⁹For European options, the put-call parity applies. As a consequence, the value of European put options written on $P(t, T + \delta)$ can be similarly calculated; see Chen (1992) for details.

$$\begin{aligned}
v(t, T, T + \delta) &= \sigma_{11}^2 \delta^2 (T - t) + (\sigma_{21}^2 + \sigma_{22}^2) \left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \\
&+ (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \left[\left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right. \\
&\quad + e^{-2\lambda \delta} \left[\frac{\delta^2 - (T + \delta - t)^2 e^{-2\lambda(T-t)}}{2\lambda} + \frac{\delta - (T + \delta - t) e^{-2\lambda(T-t)}}{2\lambda^2} + \frac{1 - e^{-2\lambda(T-t)}}{4\lambda^3} \right] \\
&\quad - \frac{1}{2\lambda} (T - t)^2 e^{-2\lambda(T-t)} - \frac{1}{2\lambda^2} (T - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{4\lambda^3} \\
&\quad - \frac{(1 - e^{-\lambda \delta}) e^{-\lambda \delta}}{\lambda^2} \left[\delta - (T + \delta - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad + \frac{1 - e^{-\lambda \delta}}{\lambda^2} \left[\frac{1 - e^{-2\lambda(T-t)}}{2\lambda} - (T - t) e^{-2\lambda(T-t)} \right] \\
&\quad + \frac{1}{\lambda} \delta e^{-\lambda \delta} \left[(T - t) e^{-2\lambda(T-t)} - \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad \left. + \frac{1}{\lambda} e^{-\lambda \delta} \left[(T - t)^2 e^{-2\lambda(T-t)} + \frac{1}{\lambda} (T - t) e^{-2\lambda(T-t)} - \frac{1 - e^{-2\lambda(T-t)}}{2\lambda^2} \right] \right] \\
&+ 2\sigma_{11}\sigma_{21}\delta(1 - e^{-\lambda \delta}) \frac{1 - e^{-\lambda(T-t)}}{\lambda^2} \\
&+ 2\sigma_{11}\sigma_{31}\delta \left[-\frac{1}{\lambda} (T - t) e^{-\lambda(T-t)} - \frac{1}{\lambda} e^{-\lambda \delta} \left(\delta - (T + \delta - t) e^{-\lambda(T-t)} \right) + 2(1 - e^{-\lambda \delta}) \frac{1 - e^{-\lambda(T-t)}}{\lambda^2} \right] \\
&+ (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \left[\left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{\lambda} \right. \\
&\quad + \frac{1}{\lambda^2} e^{-2\lambda \delta} \left[\delta - (T + \delta - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad + \frac{1}{\lambda^2} \left[- (T - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad \left. - \frac{1}{\lambda^2} e^{-\lambda \delta} \left[\delta - (2T + \delta - 2t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{\lambda} \right] \right].
\end{aligned}$$

3.3 The Shadow-Rate B-AFNS Model

We refer to the complete three-factor shadow-rate model as the B-AFNS(3) model.¹⁰ Given the above AFNS(3) shadow-rate process and the price of a shadow bond option, we are now ready to price bonds that observe the nonnegativity constraint in a B-AFNS(3) model.

Krippner (2012) provides a formula for the ZLB instantaneous forward rate, $\underline{f}(t, T)$, that applies to any Gaussian model

$$\underline{f}(t, T) = f(t, T) \Phi \left(\frac{f(t, T)}{\omega(t, T)} \right) + \omega(t, T) \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, T)}{\omega(t, T)} \right]^2 \right),$$

where $\Phi(\cdot)$ is the cumulative probability function for the standard normal distribution, $f(t, T)$ is the shadow forward rate, and $\omega(t, T)$ is related to the conditional variance appearing in the

¹⁰Following KS, the prefix “B-” refers to a shadow-rate model in the spirit of Black (1995), while the number shows the number of state variables. Krippner (2012, 2013) adopts the prefix CAB for “currency-adjusted bond.”

shadow bond option price formula as follows:

$$\omega(t, T)^2 = \frac{1}{2} \lim_{\delta \rightarrow 0} \frac{\partial^2 v(t, T, T + \delta)}{\partial \delta^2}.$$

Within the B-AFNS(3) model, the formula for the shadow forward rate, $f(t, T)$, is provided in equation (5), while after some tedious calculus, $\omega(t, T)$ takes the following form:¹¹

$$\begin{aligned} \omega(t, T)^2 = & \sigma_{11}^2(T-t) + (\sigma_{21}^2 + \sigma_{22}^2) \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \\ & + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \left[\frac{1 - e^{-2\lambda(T-t)}}{4\lambda} - \frac{1}{2}(T-t)e^{-2\lambda(T-t)} - \frac{1}{2}\lambda(T-t)^2 e^{-2\lambda(T-t)} \right] \\ & + 2\sigma_{11}\sigma_{21} \frac{1 - e^{-\lambda(T-t)}}{\lambda} + 2\sigma_{11}\sigma_{31} \left[-(T-t)e^{-\lambda(T-t)} + \frac{1 - e^{-\lambda(T-t)}}{\lambda} \right] \\ & + (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \left[-(T-t)e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right]. \end{aligned}$$

Now, the zero-coupon bond yields that observe the ZLB, denoted $\underline{y}(t, T)$, are easily calculated as

$$\underline{y}(t, T) = \frac{1}{T-t} \int_t^T \left[f(t, s) \Phi \left(\frac{f(t, s)}{\omega(t, s)} \right) + \omega(t, s) \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{f(t, s)}{\omega(t, s)} \right]^2 \right) \right] ds. \quad (6)$$

As highlighted by Krippner (2012), with Gaussian shadow-rate dynamics, the calculation of zero-coupon bond yields involves only a single integral independent of the factor dimension of the model, which greatly facilitates empirical implementation.

3.4 Market Prices of Risk

So far, the description of the B-AFNS(3) model has relied solely on the dynamics of the state variables under the Q -measure used for pricing. However, to complete the description of the model and to implement it empirically, we will need to specify the risk premiums that connect the factor dynamics under the Q -measure to the dynamics under the real-world (or historical) P -measure. It is important to note that there are no restrictions on the dynamic drift components under the empirical P -measure beyond the requirement of constant volatility. To facilitate empirical implementation, we use the extended affine risk premium developed by Cheridito et al. (2007). In the Gaussian framework, this specification implies that the risk premiums Γ_t depend on the state variables; that is,

$$\Gamma_t = \gamma^0 + \gamma^1 X_t,$$

¹¹These calculations are available from the authors upon request.

where $\gamma^0 \in \mathbf{R}^3$ and $\gamma^1 \in \mathbf{R}^{3 \times 3}$ contain unrestricted parameters.¹² The relationship between real-world yield curve dynamics under the P -measure and risk-neutral dynamics under the Q -measure is given by

$$dW_t^Q = dW_t^P + \Gamma_t dt.$$

Thus, the P -dynamics of the state variables are

$$dX_t = K^P(\theta^P - X_t)dt + \Sigma dW_t^P, \quad (7)$$

where both K^P and θ^P are allowed to vary freely relative to their counterparts under the Q -measure.

Finally, we note that the model estimation is based on the Kalman filter and described in Appendix B.

4 Data

In this section, we describe our two samples of Japanese government bond yields. The first sample is the data set examined by KS.¹³ This data set contains six maturities: six-month yields and one-, two-, four-, seven-, and ten-year yields. All yields are continuously compounded and measured weekly (Fridays) from January 6, 1995, to March 7, 2008. This “KS sample” data set is useful to study for two reasons. First, it allows a direct comparison of our two-factor B-AFNS(2) estimates with the two-factor KS estimates for an identical data sample—that is, a relatively clean “apples to apples” comparison of model results. In addition—assuming the two-factor results are comparable—we can then explore whether the KS empirical results generalize to the models with more or fewer factors. However, the KS sample ends before the recent global financial crisis episode, which was marked by extremely low bond yields in Japan and in many other countries. This recent episode is extremely interesting from a variety of economic and finance perspectives. Therefore, we also consider a second “extended KS sample” that augments the original KS sample with Japanese government zero-coupon yields downloaded from Bloomberg through November 16, 2012.¹⁴

Figure 1 shows the variation over time in four of the six yields. During 2001 through 2005 and 2009 through 2012, six-month and one-year yields are pegged near zero. These episodes are obvious candidates for possible negative shadow rates. As stressed by KS, these periods also display reduced volatility of short- and medium-term yields due to the zero bound constraint.

¹²For Gaussian models, this specification is equivalent to the essentially affine risk premium specification introduced in Duffee (2002).

¹³We thank Don Kim for sharing these data.

¹⁴During 2007 and 2008, the two spliced data series match almost exactly.

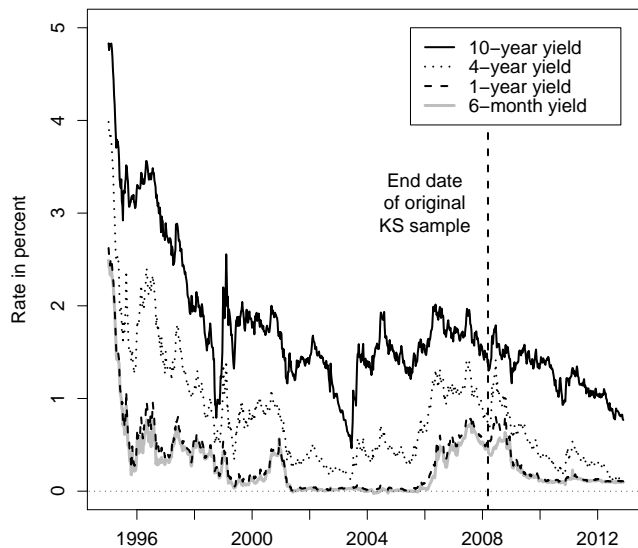


Figure 1: **Japanese Government Bond Yields.** We show time-series plots of Japanese government bond yields at weekly frequency, at maturities of 6 months, 1 year, 4 years, and 10 years. The original KS data cover the period from January 6, 1995, to March 7, 2008, and its end date is shown with a vertical dashed black line, while the extended KS sample ends on November 16, 2012.

Researchers have typically found that three factors are sufficient to model the time-variation in cross sections of bond yields (e.g., Litterman and Scheinkman, 1991). Indeed, for the KS sample, 99.86 percent of the total variation is accounted for by three factors. As Table 1 reports, the first principal component accounts for 93.7 percent of the variation in Japanese bond yields, and its loading across maturities (the associated eigenvector) is uniformly negative. Thus, like a level factor, a shock to this component changes all yields in the same direction irrespective of maturity. The second principal component accounts for 5.6 percent of the variation in these data, and like a slope factor, a shock to this component steepens or flattens the yield curve. Finally, the third component, which accounts for 0.5 percent of the variation, has a U-shaped factor loading as a function of maturity, which is naturally interpreted as a curvature factor. This pattern of level, slope, and curvature (which is also found in the extended sample) motivates our use of the Nelson-Siegel level, slope, and curvature factors for modeling Japanese bond yields, even though we emphasize that our estimated state variables are *not* identical to the principal components.

Maturity (months)	Loading on		
	First P.C.	Second P.C.	Third P.C.
6	-0.23	-0.50	0.53
12	-0.25	-0.48	0.27
24	-0.32	-0.42	-0.35
48	-0.44	-0.14	-0.58
84	-0.56	0.33	-0.08
120	-0.52	0.46	0.42
% explained	93.74	5.60	0.53

Table 1: **Factor Loadings for Japanese Government Bond Yields.** The first six rows show how bond yields at various maturities load on the first three principal components. The bottom row shows the proportion of all bond yield variability explained by each principal component. The data are weekly Japanese zero-coupon government bond yields from January 6, 1995, to March 7, 2008.

5 Results for KS Sample

In this section, we assess whether our B-AFNS implementation of Krippner’s shadow-rate framework provides good fit to the data. In particular, we assess whether a two-factor version of the B-AFNS model can closely approximate the KS B-AG2 model when estimated on the same sample of Japanese data. Furthermore, since our modeling framework can be tractably estimated with three factors, we also examine the sensitivity of the shadow rate to the number of factors in the model.

5.1 One-Factor Model Estimates

We begin the analysis by considering the simplest possible case for the shadow-rate dynamics, namely the one-factor Gaussian model of Vasiček (1977). In this one-factor case, the factor dynamics of the shadow rate s_t used for pricing under the risk-neutral Q -measure are

$$ds_t = \kappa^Q(\theta^Q - s_t)dt + \sigma dW_t^Q.$$

The instantaneous forward rate is given by

$$f(t, T) = e^{-\kappa^Q(T-t)}s_t + \theta^Q(1 - e^{-\kappa^Q(T-t)}) - \frac{1}{2}\sigma^2\left(\frac{1 - e^{-\kappa^Q(T-t)}}{\kappa^Q}\right)^2,$$

while

$$\omega(t, T)^2 = \sigma^2\frac{1 - e^{-2\kappa^Q(T-t)}}{2\kappa^Q}.$$

Allowing for time-varying risk premiums, the dynamics under the objective P -measure are fully flexible,

$$ds_t = \kappa^P(\theta^P - s_t)dt + \sigma dW_t^P.$$

Parameter	V(1)	B-V(1)
κ^P	0.0058 (0.0232)	0.0403 (0.1375)
θ^P	0.0118 (0.0103)	0.0107 (0.0169)
σ	0.0011 (0.0000)	0.0046 (0.0001)
κ^Q	0.0068 (0.0066)	0.0011 (0.0024)
θ^Q	0.9717 (0.9407)	3.7127 (7.9375)
Max log L	18,982.87	21,320.90

Table 2: **Parameter Estimates of One-Factor Models.** The estimated parameters are shown for the V(1) and B-V(1) models. The numbers in parentheses are estimated parameter standard deviations.

Errors of fitted yields				
Maturity (months)	V(1)		B-V(1)	
	Mean	RMSE	Mean	RMSE
6	8.7	10.7	2.0	4.2
12	-1.1	4.5	-0.1	0.3
24	-11.3	16.9	0.6	9.8
48	-11.9	33.5	7.3	27.7
84	10.9	53.5	10.8	48.9
120	42.5	66.2	-9.8	48.9
All yields	6.3	38.3	1.8	30.7

Table 3: **Summary Statistics of Fitted Errors in One-Factor Models.** The table presents the mean and root mean-squared error (RMSE) of the fitted bond yields from one-factor models estimated on the weekly Japanese government bond yield data over the period from January 6, 1995, to March 7, 2008. All numbers are measured in basis points.

We refer to this representation inspired by Black (1995) as the B-V(1) model. We also estimate the standard Vasiček (1977) model without the shadow-rate interpretation, denoted as the V(1) model.

Table 2 reports the estimated parameters for both one-factor models. In terms of the Q -dynamics, the very low values of κ^Q imply that the state variable is a level factor. This is also reflected in its very high persistence under the P -dynamics. The estimated mean values θ^P , which are the average levels of the state variable, are about the same in each model. The largest difference between the models is that the B-V(1) model has an estimated factor volatility about four times larger than in the V(1) model.

Table 3 reports the summary statistics of the fitted errors for the V(1) and B-V(1) models. The better fit of the B-V(1) model across all yield maturities is notable, with an average root mean-squared error (RMSE) improvement of 7.6 basis points. This better fit can also be seen in the higher likelihood value of the B-V(1) model.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$
$K_{1,\cdot}^P$	-0.3555 (2.5725)	-0.3919 (2.6886)	0.0555 (0.2343)	$\Sigma_{1,\cdot}$	0.0306 (0.0041)	0
$K_{2,\cdot}^P$	0.6184 (2.6747)	0.5679 (2.7882)	-0.0362 (0.3584)	$\Sigma_{2,\cdot}$	-0.0311 (0.0042)	0.0032 (0.0001)

Table 4: **Parameter Estimates of the AFNS(2) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the AFNS(2) model. The associated estimated λ is 0.0401 (0.0059) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 22,973.19.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$
$K_{1,\cdot}^P$	0.4688 (0.2516)	0.6331 (0.2783)	0.1029 (0.0388)	$\Sigma_{1,\cdot}$	0.0081 (0.0003)	0
$K_{2,\cdot}^P$	-0.2517 (0.2605)	-0.3230 (0.3130)	-0.0951 (0.0311)	$\Sigma_{2,\cdot}$	-0.0072 (0.0003)	0.0057 (0.0002)

Table 5: **Parameter Estimates of the B-AFNS(2) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the B-AFNS(2) model. The associated estimated λ is 0.1448 (0.0040) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 23,786.44.

5.2 Two-Factor Model Estimates

To most closely approximate the KS B-AG2 model, we estimate a two-factor version of the B-AFNS model that has level and slope factors but no curvature factor. This model is characterized by a shadow rate given by

$$s_t = X_t^1 + X_t^2.$$

Its dynamics of the state variables (X_t^1, X_t^2) used for pricing under the risk-neutral Q -measure have the following structure:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = - \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dX_t^{2,Q} \end{pmatrix}.$$

As for the P -dynamics, we focus on the most flexible specification with full K^P matrix

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P \\ \kappa_{21}^P & \kappa_{22}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \end{pmatrix}.$$

This model has a total of ten parameters, two less than the canonical B-AG2 model used by KS.

We estimate both the regular version of the model above without any constraints related to

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	θ^P	K^Q	$K_{\cdot,1}^Q$	$K_{\cdot,2}^Q$	θ^Q		Σ
$K_{1,\cdot}^P$	0.0204 (0.0198)	0	0	$K_{1,\cdot}^Q$	0.1669 (0.0394)	-0.5275 (0.0641)	-0.0191 (0.0115)	$\Sigma_{1,1}$	0.0035 (0.0001)
$K_{2,\cdot}^P$	-0.1617 (0.1724)	0.2664 (0.2765)	0	$K_{2,\cdot}^Q$	0.1771 (0.1654)	-0.2320 (0.2795)	-0.0163 (0.0077)	$\Sigma_{2,2}$	0.0026 (0.0001)

Table 6: **Parameter Estimates of the B-AG2 Model in KS.** The estimated parameters of the K^P and K^Q matrices, the θ^P and θ^Q vectors, and the Σ matrix are shown for the B-AG2 model as reported by KS. The constant ρ in the specification of s_t is 0.0381 (0.0008). The numbers in parentheses are estimated parameter standard deviations.

Errors of fitted yields					
Maturity (months)	AFNS(2)		B-AFNS(2)		B-AG2
	Mean	RMSE	Mean	RMSE	$\widehat{\sigma}_\varepsilon(\tau)$
6	4.8	6.0	5.0	6.7	5
12	0.0	0.0	-0.1	0.4	0
24	-2.6	9.3	-4.3	8.1	7
48	2.2	17.2	-2.8	11.1	8
84	9.8	21.9	4.5	15.0	2
120	0.0	0.0	-0.1	3.1	12
All yields	2.4	12.2	0.4	8.8	n.a.

Table 7: **Summary Statistics of Fitted Errors in Two-Factor Models.** The table presents the mean and root mean-squared error of the fitted bond yields from two-factor models estimated on the weekly Japanese government bond yield data over the period from January 6, 1995, to March 7, 2008. Also shown are the estimated error standard deviations reported by KS for their B-AG2 model. All numbers are measured in basis points.

the ZLB, denoted as the AFNS(2) model, and the corresponding shadow-rate model, denoted as the B-AFNS(2) model. Tables 4 and 5 report the estimated parameters for the AFNS(2) and B-AFNS(2) models, respectively. For comparison, Table 6 also shows the parameters of the B-AG2 model as estimated by KS, although it is difficult to directly compare the B-AFNS and B-AG2 model parameters due to the latent nature of the yield factors and differences in their econometric identification.

In the AFNS(2) and B-AFNS(2) models, the estimated λ values are relatively close. Beyond that, the estimated mean-reversion matrix, mean vector, and volatility matrix share only a few broad similarities such as positive θ_1^P , negative θ_2^P , and negative σ_{21} parameters, but in terms of magnitudes the differences are sizeable. Notably, relative to the B-AG2 model, the factors of the AFNS(2) and B-AFNS(2) models are more volatile and less persistent.

Table 7 reports summary statistics for the fit of the two-factor models. The AFNS(2) model performs reasonably well, but the B-AFNS(2) model has smaller yield RMSEs. The fit of the B-AFNS(2) model is comparable to the B-AG2 model estimated in KS even though

Errors of fitted yields				
Maturity (months)	AFNS(3)		B-AFNS(3)	
	Mean	RMSE	Mean	RMSE
6	0.0	0.0	-0.1	0.6
12	0.0	2.0	-0.1	0.3
24	0.0	0.3	0.5	3.5
48	-0.9	4.0	-0.6	4.5
84	0.0	0.0	-0.1	0.9
120	-3.0	21.8	-3.2	16.1
All yields	-0.7	9.1	-0.6	7.0

Table 8: **Summary Statistics of Fitted Errors in Three-Factor Models.** The table presents the mean and root mean-squared error of the fitted bond yields from three-factor models estimated on the weekly Japanese government bond yield data over the period from January 6, 1995, to March 7, 2008. All numbers are measured in basis points.

the B-AFNS(2) model has fewer parameters under the Q -dynamics used for pricing.¹⁵

5.3 Three-Factor Model Estimates

In this section, we extend the analysis to assess whether the KS results generalize to three-factor models. We compare the B-AFNS(3) model to the unconstrained AFNS(3) model when both are estimated on the KS sample of Japanese data.

In the AFNS(3) model, the risk-neutral Q -dynamics used for pricing are as detailed in Section 3, while we assume fully flexible factor dynamics under the P -measure as in the previous section:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

Table 8 reports the summary statistics of the fitted errors of the regular AFNS(3) model as well as its shadow-rate version. Similar to what we observed for the two-factor models, the shadow-rate model clearly outperforms its standard counterpart when it comes to model fit. However, in comparing model fit across the two- and three-factor models, we note that the AFNS(3) model outperforms the B-AFNS(2) model, while the B-AFNS(3) model outperforms both of them.

Tables 9 and 10 contain the estimated parameters for the AFNS(3) and B-AFNS(3) models. With the exception of the estimated λ values and Σ volatility matrices, there are large differences in both signs and magnitudes for most parameters across the two models. Fur-

¹⁵We do not report our estimated error standard deviations, $\hat{\sigma}_\varepsilon(\tau_i)$, as they are typically very close to the RMSEs. By implication, the RMSEs shown in Table 7 are comparable to the $\hat{\sigma}_\varepsilon(\tau_i)$ reported by KS.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	$K_{\cdot,3}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$	$\Sigma_{\cdot,3}$
$K_{1,\cdot}^P$	2.3098 (1.3248)	2.9063 (1.5991)	-0.8699 (0.4927)	0.0523 (0.0952)	$\Sigma_{1,\cdot}$	0.0136 (0.0006)	0	0
$K_{2,\cdot}^P$	-1.1272 (1.2990)	-1.3481 (1.5670)	0.3993 (0.4842)	-0.0458 (0.0731)	$\Sigma_{2,\cdot}$	-0.0132 (0.0006)	0.0027 (0.0001)	0
$K_{3,\cdot}^P$	1.4331 (1.7396)	2.0048 (2.1814)	0.4314 (0.7018)	-0.0310 (0.0088)	$\Sigma_{3,\cdot}$	-0.0183 (0.0012)	-0.0001 (0.0006)	0.0153 (0.0004)

Table 9: **Parameter Estimates of the AFNS(3) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the AFNS(3) model. The associated estimated λ is 0.4157 (0.0043) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 25,547.96.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	$K_{\cdot,3}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$	$\Sigma_{\cdot,3}$
$K_{1,\cdot}^P$	0.3283 (0.5718)	0.6803 (0.6158)	-0.2379 (0.2234)	0.0034 (0.0829)	$\Sigma_{1,\cdot}$	0.0142 (0.0006)	0	0
$K_{2,\cdot}^P$	0.7081 (0.5359)	0.9228 (0.5954)	-0.1726 (0.2202)	0.0269 (0.1813)	$\Sigma_{2,\cdot}$	-0.0131 (0.0005)	0.0048 (0.0002)	0
$K_{3,\cdot}^P$	1.5256 (0.7108)	1.5075 (0.7812)	-0.2517 (0.2390)	0.1194 (0.5139)	$\Sigma_{3,\cdot}$	-0.0133 (0.0010)	0.0057 (0.0009)	0.0127 (0.0005)

Table 10: **Parameter Estimates of the B-AFNS(3) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the B-AFNS(3) model. The associated estimated λ is 0.5101 (0.0035) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 26,334.34.

thermore, the estimated parameters for the level and slope factors in the AFNS(3) models only vaguely resemble the corresponding parameters in the AFNS(2) models, but this is a common feature when estimating flexible latent factor models such as ours.¹⁶

5.4 Shadow Rate Comparisons Across Models

Figure 2 shows the instantaneous shadow short-rate paths implied by our one-, two-, and three-factor shadow-rate models. Also, for comparison, we include the shadow-rate path from the B-AG2 model as estimated by KS. As shown in Table 11, which contains the pairwise correlations between the estimated shadow-rate paths, the correlations across models range from 0.887 to 0.993, the latter being the correlation between the results from the B-AFNS(2) and B-AG2 models that could be expected to be high due to the identical number of state variables. Thus, the models tend to agree on the general variation in the shadow rate. However, when the shadow rate is negative, there can be pronounced differences among the estimated shadow rates of various models. The shadow rate from the B-AFNS(3) model

¹⁶This is part of the reason why CDR recommend focusing on parsimonious specifications of the AFNS models, say, with a diagonal Σ matrix and additional restrictions on K^P as in Christensen, Lopez, and Rudebusch (2010).

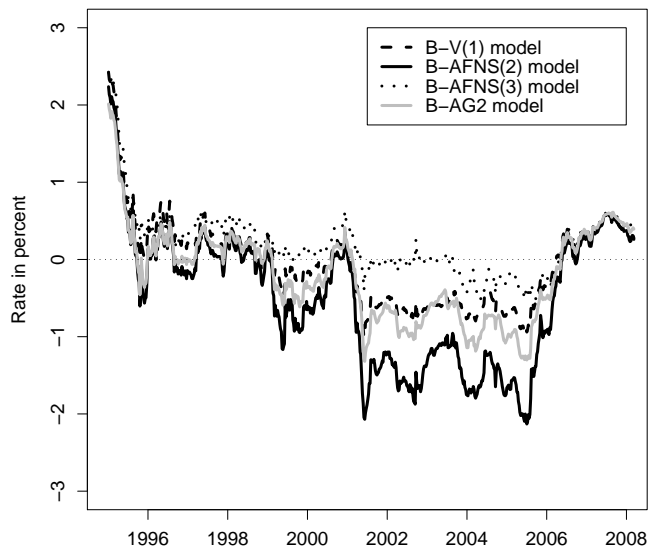


Figure 2: **Model-Implied Shadow Rates.** Illustration of the model-implied shadow rate from the B-V(1), B-AFNS(2), and B-AFNS(3) models. For comparison, we include the shadow rate from the B-AG2 model estimated by KS.

Correlation	B-V(1)	B-AFNS(2)	B-AFNS(3)	B-AG2
B-V(1)	1	0.973	0.936	0.974
B-AFNS(2)		1	0.887	0.993
B-AFNS(3)			1	0.914
B-AG2				1

Table 11: **Pairwise Correlations of Shadow-Rate Paths.** The table contains the pairwise correlations between the shadow-rate paths from the B-V(1), B-AFNS(2), B-AFNS(3), and B-AG2 models. The estimations are based on the KS sample.

is generally the least negative. The shadow rates from the B-V(1) and B-AG2 models are more negative and often quite close to each other. The B-AFNS(2) model provides the most negative shadow rates.

These shadow-rate results can be usefully compared to two other studies. Ueno et al. (2006) use Japanese data for a shorter sample (2003-2006) and a different maturity composition (ten maturities ranging from six months to twenty years) than we do. In a calibrated one-factor version of the Black (1995) model, they obtain a shadow rate lower than -5 percent for most of their sample, with the lowest reading falling below -15 percent in the summer of 2002.¹⁷ On the other hand, Ichiue and Ueno (2007), who use the Kalman filter to estimate a

¹⁷Krippner (2012) uses a two-factor shadow-rate model applied to U.S. Treasury data and also reports large

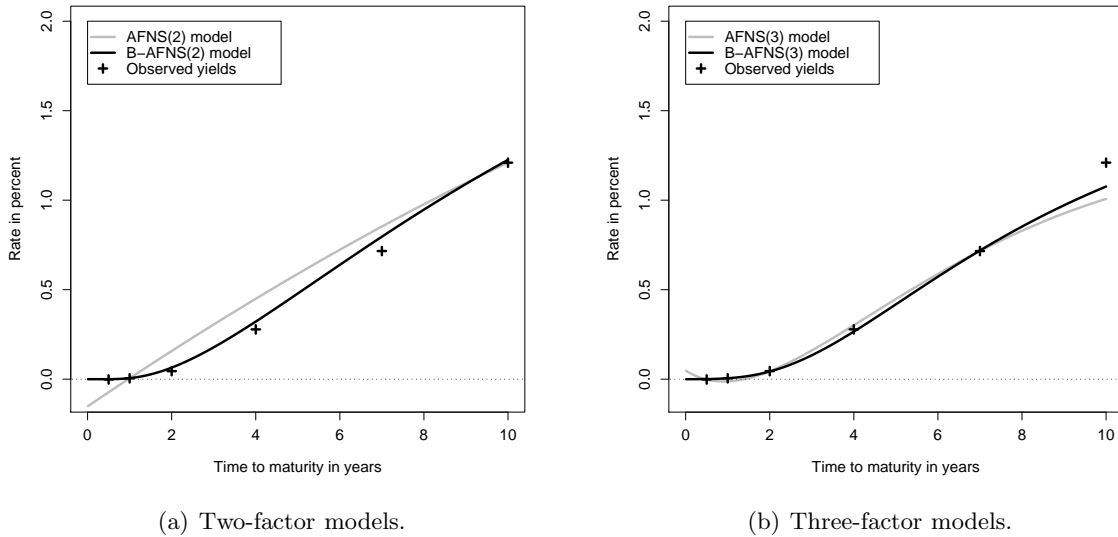


Figure 3: Fitted Yield Curves on July 1, 2005. The figure to the left illustrates the fitted yield curves from the AFNS(2) and B-AFNS(2) models on July 1, 2005. Also shown are the six observed yields on that date. The figure to the right shows the corresponding results for the AFNS(3) and B-AFNS(3) models.

two-factor shadow-rate model on monthly Japanese government bond yields from July 1996 to March 2006, report shadow-rate values in a range from -1 to -0.5 percent for the 2001-2005 period, with time-series variation very similar to our results. Combined, these results suggest that it appears to be important to perform a full empirical implementation based on maximum likelihood that simultaneously accounts for the time-series and cross-sectional variation in the data.

To further illustrate differences between the standard and shadow-rate models, we examine the two- and three-factor models' implications on a specific date, July 1, 2005, when the shadow rate attains its lowest value according to most models shown in Figure 2. Figure 3(a) illustrates observed yields on this date as well as fitted yield curves from the AFNS(2) and B-AFNS(2) models, while Figure 3(b) shows the corresponding output for the AFNS(3) and B-AFNS(3) models. For the two-factor models, we note that the AFNS(2) model has difficulty matching the kink in the observed yields around the two-year maturity point, which is very pronounced in the data for this period. On the other hand, for the three-factor models this distinction between standard and shadow-rate models is much less clear. It appears as if the plain-vanilla AFNS(3) model has sufficient flexibility to handle the kink even on this very challenging day in the sample. This also explains the relatively small differences in the fitted errors across these two models.

Figure 4(a) shows the short rate projections on July 1, 2005, from the AFNS(2) and B-

negative shadow rates that average about -4 percent since August 2010.

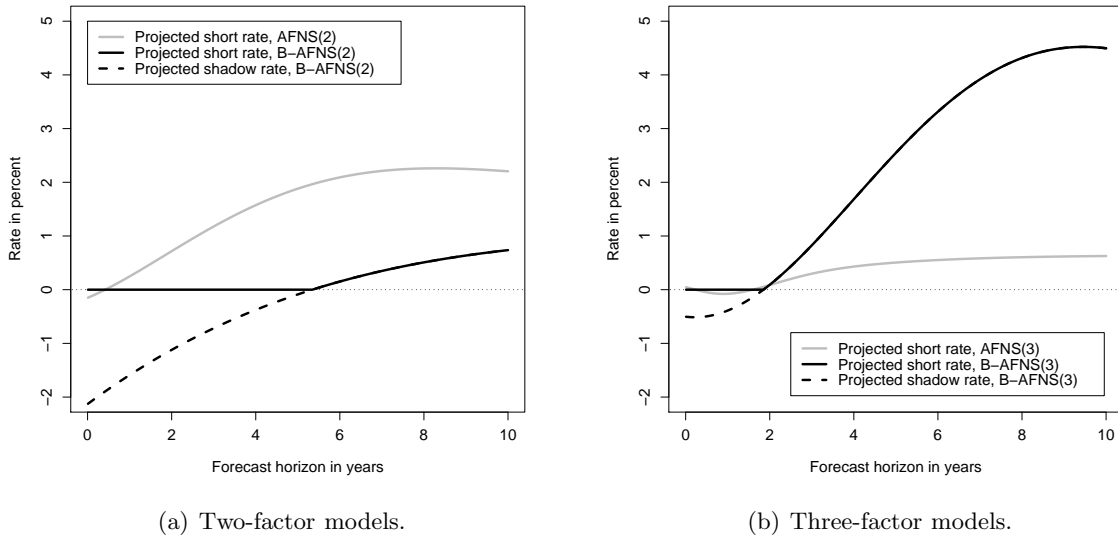


Figure 4: **Shadow-Rate Projections on July 1, 2005.** The figure to the left illustrates the short rate projections from the AFNS(2) and B-AFNS(2) models along with the shadow-rate projection from the latter model on July 1, 2005. The figure to the right shows the corresponding results for the AFNS(3) and B-AFNS(3) models.

AFNS(2) models, while Figure 4(b) shows the corresponding projections from the AFNS(3) and B-AFNS(3) models. Also shown in the figures are the shadow-rate projections. Note that the estimated shadow rate in the B-AFNS(2) model is notably more negative than in the B-AFNS(3) model on this date. Furthermore, with more persistent factor dynamics, the B-AFNS(2) reversal to steady state is very gradual. Short rates would remain near zero for another five years according to the B-AFNS(2) model, as opposed to about two years according to the B-AFNS(3) model. For the period after liftoff, the difference in persistence leads to an even larger divergence in projections, as the B-AFNS(3) model suggests that the policy rate would go above 2 percent in just four years, while liftoff is well beyond even a ten-year forecast horizon according to the two-factor model.

The low persistence of the B-AFNS(3) model is an issue that frequently plagues flexible models and is tied to the length and nature of the data sample. Because interest rates are highly persistent, empirical autoregressive models, including dynamic term structure models like ours, suffer from potentially substantial finite-sample estimation bias. Bauer, Rudebusch, and Wu (2012) provide a complete discussion of the finite-sample bias in empirical affine Gaussian term structure models and simulation-based methods to eliminate it. Furthermore, Christensen and Gillan (2012) present a simplification of their approach that can be used for latent-factor Kalman filter estimations like the ones in this paper. Alternatively, higher persistence can be imposed by assuming a unit-root property for the most persistent factor;

Par.	B-V(1)
κ^P	0.0222 (0.1481)
θ^P	0.0100 (0.0300)
σ	0.0042 (0.0001)
κ^Q	0.0003 (0.0001)
θ^Q	12.73 (5.9381)
Max log L	28,566.74

Table 12: **Estimated Parameters of One-Factor Model.** The estimated parameters are shown for the B-V(1) model using the extended KS sample. The numbers in parentheses are estimated parameter standard deviations.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$
$K_{1,\cdot}^P$	0.3430 (0.2093)	0.4732 (0.2364)	0.1026 (0.0718)	$\Sigma_{1,\cdot}$	0.0076 (0.0003)	0
$K_{2,\cdot}^P$	-0.1381 (0.1955)	-0.1756 (0.2428)	-0.0927 (0.0497)	$\Sigma_{2,\cdot}$	-0.0070 (0.0003)	0.0048 (0.0001)

Table 13: **Estimated Parameters of the B-AFNS(2) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the B-AFNS(2) model using the extended KS sample. The associated estimated λ is 0.1260 (0.0039) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 31,987.75.

K^P	$K_{\cdot,1}^P$	$K_{\cdot,2}^P$	$K_{\cdot,3}^P$	θ^P	Σ	$\Sigma_{\cdot,1}$	$\Sigma_{\cdot,2}$	$\Sigma_{\cdot,3}$
$K_{1,\cdot}^P$	1.0366 (0.9525)	1.4585 (1.3468)	-0.4774 (0.3545)	0.0031 (0.1122)	$\Sigma_{1,\cdot}$	0.0208 (0.0006)	0	0
$K_{2,\cdot}^P$	0.1826 (0.9297)	0.3294 (1.2624)	-0.0575 (0.3330)	0.0346 (0.2642)	$\Sigma_{2,\cdot}$	-0.0190 (0.0006)	0.0040 (0.0001)	0
$K_{3,\cdot}^P$	-0.4283 (1.2960)	-1.1078 (1.7500)	0.4094 (0.4996)	0.1207 (0.6260)	$\Sigma_{3,\cdot}$	-0.0288 (0.0009)	-0.0010 (0.0006)	0.0181 (0.0004)

Table 14: **Estimated Parameters of the B-AFNS(3) Model.** The estimated parameters of the K^P matrix, the θ^P vector, and the Σ matrix are shown for the B-AFNS(3) model using the extended KS sample. The associated estimated λ is 0.4893 (0.0044) with maturity measured in years. The numbers in parentheses are estimated parameter standard deviations. The maximum log likelihood value is 35,568.52.

see Christensen and Rudebusch (2012) for an example. Dealing with such estimation issues is outside the scope of this paper but is an important issue for future applied work.

6 Results for Extended KS Sample

In this section, we estimate all three shadow-rate models on the extended KS sample that ends on November 16, 2012. For a start, we compare the estimated parameters from the extended

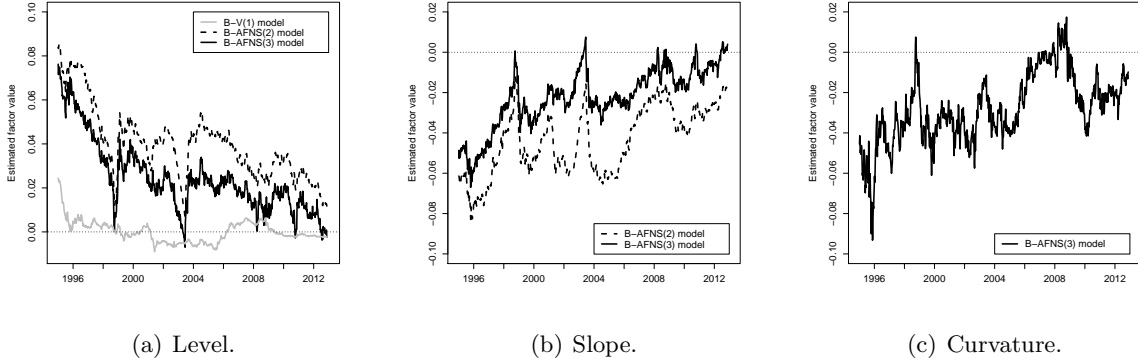


Figure 5: **Estimated Factors in Shadow-Rate Models.** Illustration of the estimated factors in the B-V(1), B-AFNS(2), and B-AFNS(3) models from the extended KS sample.

KS sample to those reported in the previous section for the original KS sample. Table 12 reports the estimated parameters for the B-V(1) model. Relative to the results reported for the B-V(1) model in Table 2, there is an increase in persistence and a decline in volatility, likely because the added four years of data have lower and more stable yields than the prior 13 years of the sample, as can be seen in Figure 1. When we compare the results for the B-AFNS(2) model on the extended sample reported in Table 13 to the corresponding results for the KS sample in Table 5, we see a similar pattern of increased persistence and lower volatility for both the level and the slope factor with only minor changes to their estimated mean levels. For the B-AFNS(3) model, whose estimated parameters for the extended sample are shown in Table 14, a slightly different pattern is observed, with the level factor becoming less persistent and more volatile, while only the slope and curvature factors follow the pattern of greater persistence observed in the more parsimonious models.

Next, we compare the estimated factors across models. Figure 5(a) shows the estimated level factors from all three models. There is a very high correlation (96.1 percent) between the level factors in the B-AFNS(2) and B-AFNS(3) models, whereas the connection to the factor in the B-V(1) model is much weaker. Figure 5(b) shows the estimated slope factors from the B-AFNS(2) and B-AFNS(3) models, where we again see a high correlation of 88.5 percent. Furthermore, within each model, there are interesting relationships. In the B-AFNS(2) model, the correlation between its factors is -90.5 percent, while the corresponding correlation between level and slope in the B-AFNS(3) model is as high as -97.5 percent. Also, its correlation is -73.2 percent between level and curvature and 77.2 percent between slope and curvature. Thus, the state variables in the AFNS models are highly correlated, which supports the use of flexible specifications for this data.

Figure 6 shows the estimated shadow-rate paths implied by the three models on the extended sample. As in Figure 2, there is much less disagreement across models when the

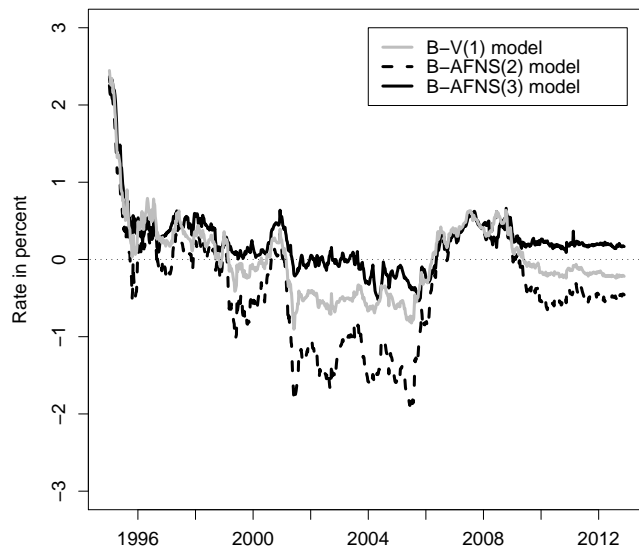


Figure 6: **Model-Implied Shadow Rates from the Extended KS Sample.** Illustration of the model-implied shadow rate from the B-V(1), B-AFNS(2), and B-AFNS(3) models.

instantaneous rate is in positive territory (1995-1998 and 2006-2009), presumably because short-term yields are far enough away from the ZLB that the fit to the cross section of yields requires a positive short rate and all models are forced to accommodate that. On the other hand, when all models indicate that the shadow rate is in negative territory—the 2001-2005 period is a particularly strong example of this—there is notable disagreement about its exact level across the one-, two-, and three-factor models. Unreported results show that, even within each model class, there can be disagreement across specifications about how negative the shadow rate is during this important part of the sample. Thus, great care must be taken in interpreting the magnitude of the negative shadow rate as a measure of the stance of monetary policy when yields are near the ZLB, as suggested by Krippner (2013). At a minimum, a number of model specifications should be analyzed to verify the robustness of the results.

An alternative way of looking at these results is to note that the shadow-rate paths are very highly correlated across models. As a consequence, all the models would tend to agree in terms of qualitative statements about the movements in the shadow rate from one period to another. This kind of inference seems to be more robust. Building on this insight, we find it safe to say that the shadow rate in Japan has declined over the extended part of our sample since March 2008 and is lower at the end of the sample than it was back then. However,

whether it is currently negative and, if so, to what extent is not really possible to tell. A much more detailed analysis of each model class would be required, including a validation of their performance along multiple dimensions such as their ability to replicate the yield volatility and excess returns seen in the data, as illustrated in KS.

7 Conclusion

To adapt the Gaussian term structure model to the recent near-zero interest rate environment, we have combined the arbitrage-free Nelson-Siegel model dynamics with the shadow-rate framework of Krippner (2012). To investigate the suitability of this pairing, we estimate variants of this model using near-zero Japanese yields. We find that model fit and the magnitude of estimated negative shadow rates appear to be sensitive to the number of factors included in the estimation. Other aspects of model specification—such as the sample maturity selection or sample period (including the ratio of near-ZLB yields to normal yield observations)—would also likely have an important influence on the shadow rate.

As emphasized by Krippner (2013), one potential use for shadow-rate models is to gauge the stance of monetary policy in near-zero yield environments where the target rate itself is constrained at its effective lower boundary and, by implication, is *not* a useful indicator of the stance of monetary policy. Regarding this type of inference, our results suggest fairly robust qualitative or directional, rather than quantitative, statements can be made. For gauging the current and future stance of monetary policy, the B-AFNS(3) model introduced in this paper appears competitive. Still, care must be taken to assess the sensitivity of the results to the number of factors used in estimation and other modeling choices. Finding a preferred shadow-rate specification of the factor dynamics and dealing with any finite-sample estimation bias is left to future research.

Appendix A: Bond Option Pricing in the AFNS(3) Model

In this appendix we derive the value of a European call option with maturity at time T and strike K written on the zero-coupon bond with maturity $T + \delta$ when it is assumed that the state variables have the AFNS(3) Q -dynamics in equation (4).

In short form, the factor Q -dynamics in the AFNS(3) model are

$$dX_t = K^Q(\theta^Q - X_t)dt + \Sigma dW_t^Q,$$

while the instantaneous risk-free rate is $r_t = \rho_1' X_t$.

Now, recall that the value of the zero-coupon bond that matures at $T + \delta$ is

$$P(t, T + \delta) = \exp(A(t, T + \delta) + B(t, T + \delta)' X_t),$$

where $A(t, T + \delta)$ and $B(t, T + \delta)$ are the unique solutions to the following ordinary differential equations (ODE) as in Duffie and Kan (1996)

$$\begin{aligned} \frac{dB(t, T + \delta)}{dt} &= \rho_1 + (K^Q)' B(t, T + \delta), \quad B(T + \delta, T + \delta) = 0, \\ \frac{dA(t, T + \delta)}{dt} &= -\frac{1}{2} \sum_{j=1}^n (\Sigma' B(t, T + \delta) B(t, T + \delta)' \Sigma)_{j,j}, \quad A(T + \delta, T + \delta) = 0. \end{aligned}$$

By Ito's lemma, the Q -dynamics of $P(t, T + \delta)$ are

$$\begin{aligned} dP(t, T + \delta) &= P(t, T + \delta) \left[\frac{dA(t, T + \delta)}{dt} + \frac{dB(t, T + \delta)' X_t}{dt} \right] dt + P(t, T + \delta) B(t, T + \delta)' dX_t \\ &\quad + \frac{1}{2} P(t, T + \delta) dX_t' B(t, T + \delta) B(t, T + \delta)' dX_t \\ &= P(t, T + \delta) \left[-\frac{1}{2} \sum_{j=1}^3 (\Sigma' B(t, T + \delta) B(t, T + \delta)' \Sigma)_{j,j} + (\rho_1 + (K^Q)' B(t, T + \delta))' X_t \right] dt \\ &\quad + P(t, T + \delta) B(t, T + \delta)' [-\kappa^Q X_t dt + \Sigma dW_t^Q] \\ &\quad + \frac{1}{2} P(t, T + \delta) \sum_{j=1}^3 (\Sigma' B(t, T + \delta) B(t, T + \delta)' \Sigma)_{j,j} dt \\ &= \rho_1' X_t P(t, T + \delta) dt + P(t, T + \delta) B(t, T + \delta)' \Sigma dW_t^Q. \end{aligned}$$

Since $\rho_1' X_t = r_t$, this reduces to

$$dP(t, T + \delta) = r_t P(t, T + \delta) dt + P(t, T + \delta) B(t, T + \delta)' \Sigma dW_t^Q. \quad (8)$$

These are the bond price dynamics under the risk-neutral measure where the riskless asset has been used as the deflator and foundation for the martingale measure applied for asset pricing.

The Forward Measure

Now, an alternative martingale measure turns out to be convenient for asset pricing for the problem at hand. This measure is frequently referred to as the forward measure and uses the zero-coupon bond price with the same maturity as the option, that is $P(t, T)$, as deflator instead of the riskless asset.

To begin, let $Z(t, T, T + \delta)$ denote the zero-coupon bond price underlying the option deflated by the zero-coupon bond $P(t, T)$

$$Z(t, T, T + \delta) = \frac{P(t, T + \delta)}{P(t, T)}.$$

By Ito's lemma,

$$\begin{aligned}
dZ(t, T, T + \delta) &= \frac{1}{P(t, T)} dP(t, T + \delta) - \frac{P(t, T + \delta)}{P(t, T)} dP(t, T) \\
&+ \frac{1}{2} \begin{pmatrix} dP(t, T + \delta) & dP(t, T) \end{pmatrix} \begin{pmatrix} 0 & \frac{-1}{P(t, T)^2} \\ \frac{-1}{P(t, T)^2} & \frac{2P(t, T + \delta)}{P(t, T)^3} \end{pmatrix} \begin{pmatrix} dP(t, T + \delta) \\ dP(t, T) \end{pmatrix}.
\end{aligned}$$

Using the result in equation (8), this reduces to

$$\begin{aligned}
dZ(t, T, T + \delta) &= r_t Z(t, T, T + \delta) dt + Z(t, T, T + \delta) B(t, T + \delta)' \Sigma dW_t^Q \\
&- r_t Z(t, T, T + \delta) dt - Z(t, T, T + \delta) B(t, T)' \Sigma dW_t^Q \\
&- \frac{1}{P(t, T)^2} dP(t, T + \delta) dP(t, T) + \frac{P(t, T + \delta)}{P(t, T)^3} dP(t, T)^2 \\
&= Z(t, T, T + \delta) [B(t, T + \delta) - B(t, T)]' \Sigma dW_t^Q \\
&- Z(t, T, T + \delta) \sum_{j=1}^3 (\Sigma' B(t, T + \delta) B(t, T)' \Sigma)_{j,j} dt \\
&+ Z(t, T, T + \delta) \sum_{j=1}^3 (\Sigma' B(t, T) B(t, T)' \Sigma)_{j,j} dt.
\end{aligned}$$

We can now define the new measure by determining the Girsanov transformation, which is the process $g(t, T)$ that shows the change in drift from the old measure to the new measure and establishes the connection between the old Brownian motion and the new Brownian motion

$$dW_t^{Q^T} = dW_t^Q - g(t, T) dt.$$

Inserting this in the dynamics above, it follows

$$\begin{aligned}
dZ(t, T, T + \delta) &= Z(t, T, T + \delta) [B(t, T + \delta) - B(t, T)]' \Sigma [dW_t^{Q^T} + g(t, T) dt] \\
&- Z(t, T, T + \delta) \sum_{j=1}^3 (\Sigma' B(t, T + \delta) B(t, T)' \Sigma)_{j,j} dt \\
&+ Z(t, T, T + \delta) \sum_{j=1}^3 (\Sigma' B(t, T) B(t, T)' \Sigma)_{j,j} dt.
\end{aligned}$$

Since the new measure should be a martingale measure that can be used for pricing, $g(t, T)$ is chosen such that the drift in the dynamics above is eliminated

$$-\sum_{j=1}^3 \left(\Sigma' [B(t, T + \delta) - B(t, T)] B(t, T)' \Sigma \right)_{j,j} + [B(t, T + \delta) - B(t, T)]' \Sigma g(t, T) = 0 \quad \text{for } t \in [0, T]. \quad (9)$$

Thus, under the forward Q^T -measure, it holds that

$$dZ(t, T, T + \delta) = Z(t, T, T + \delta) [B(t, T + \delta) - B(t, T)]' \Sigma dW_t^{Q^T}.$$

Option Pricing under the Forward Measure

Now, the key thing is the dynamics of the deflated zero-coupon bond price underlying the option, i.e. $P(t, T + \delta)$, under the T -forward measure

$$dZ(t, T, T + \delta) = Z(t, T, T + \delta) [B(t, T + \delta) - B(t, T)]' \Sigma dW_t^{Q^T}.$$

In integral form, this converts into

$$Z(s, T, T + \delta) = Z(t, T, T + \delta) + \int_t^s Z(u, T, T + \delta) [B(u, T + \delta) - B(u, T)]' \Sigma dW_u^{Q^T}, \quad s \in (t, T).$$

Due to the martingale property of the Ito integral, it follows that

$$E_t^{Q^T} [Z(s, T, T + \delta)] = Z(t, T, T + \delta).$$

Now, we focus on pricing bond options under the T -forward measure. To begin, consider the call option with maturity at T and strike price K written on the zero-coupon bond maturing at $T + \delta$. Denote its price by $C(t, T, T + \delta; K)$. Due to the Q^T -martingale property of the deflated bond price dynamics, it holds that

$$\frac{C(t, T, T + \delta; K)}{P(t, T)} = E_t^{Q^T} \left[\frac{C(T, T, T + \delta; K)}{P(T, T)} \right].$$

However, at maturity T , $P(T, T) = 1$ and $C(T, T, T + \delta; K) = \max\{P(T, T + \delta) - K, 0\}$. This implies that the call option price can be calculated as

$$C(t, T, T + \delta; K) = P(t, T) E_t^{Q^T} \left[P(T, T + \delta) \mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right] - KP(t, T) E_t^{Q^T} \left[\mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right].$$

To calculate these two contingent expectations, we exploit the properties of the $Z(t, T, T + \delta)$ process. At time T , it holds that

$$Z(T, T, T + \delta) = \frac{P(T, T + \delta)}{P(T, T)} = P(T, T + \delta).$$

Thus, the states of the world where $P(T, T + \delta)$ are above the strike K are the states of the world where

$$Z(T, T, T + \delta) \geq K.$$

Since $Z(t, T, T + \delta)$ is a log-normal process, we take its log

$$Y(t, T, T + \delta) = \ln Z(t, T, T + \delta).$$

By Ito's lemma, it holds that

$$\begin{aligned} dY(t, T, T + \delta) &= \frac{1}{Z(t, T, T + \delta)} dZ(t, T, T + \delta) - \frac{1}{2} \frac{1}{Z(t, T, T + \delta)^2} (dZ(t, T, T + \delta))^2 \\ &= [B(t, T + \delta) - B(t, T)]' \Sigma dW_t^{Q^T} \\ &\quad - \frac{1}{2} \sum_{j=1}^3 (\Sigma' [B(t, T + \delta) - B(t, T)] [B(t, T + \delta) - B(t, T)]' \Sigma)_{j,j} dt. \end{aligned}$$

In integral form, this converts into

$$\begin{aligned} Y(T, T, T + \delta) &= Y(t, T, T + \delta) - \frac{1}{2} \int_t^T \sum_{j=1}^3 (\Sigma' [B(s, T + \delta) - B(s, T)] [B(s, T + \delta) - B(s, T)]' \Sigma)_{j,j} ds \\ &\quad + \int_t^T [B(s, T + \delta) - B(s, T)]' \Sigma dW_s^{Q^T}. \end{aligned}$$

It follows that $Y(T, T, T + \delta)$ is normally distributed

$$Y(T, T, T + \delta) \sim N(m_Y(t, T, T + \delta), v_Y(t, T, T + \delta)),$$

where $m_Y(t, T, T + \delta)$ and $v_Y(t, T, T + \delta)$ will be determined below.

Now, the call option is in the money whenever

$$Y(T, T, T + \delta) = m_Y(t, T, T + \delta) + \sqrt{v_Y(t, T, T + \delta)}X_Y \geq \ln K,$$

where X_Y is a standard normally distributed variable. Equivalently,

$$X_Y \geq \frac{\ln K - m_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}.$$

It follows that the second part of the option payment can be calculated as

$$\begin{aligned} C_2(t, T, T + \delta; K) &= -KP(t, T)E_t^{Q^T} \left[\mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right] \\ &= -KP(t, T) \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln K - m_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}}^{\infty} e^{-\frac{1}{2}X_Y^2} dX_Y \\ &= -KP(t, T) \Phi \left(\frac{m_Y(t, T, T + \delta) - \ln K}{\sqrt{v_Y(t, T, T + \delta)}} \right). \end{aligned}$$

As for the first part of the option payment, it holds that

$$\begin{aligned} C_1(t, T, T + \delta; K) &= P(t, T)E_t^{Q^T} \left[P(T, T + \delta) \mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right] \\ &= P(t, T)E_t^{Q^T} \left[Z(T, T, T + \delta) \mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right] \\ &= P(t, T)E_t^{Q^T} \left[e^{Y(T, T, T + \delta)} \mathbf{1}_{\{P(T, T + \delta) \geq K\}} \right] \\ &= P(t, T) \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln K - m_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}}^{\infty} e^{m_Y(t, T, T + \delta) + \sqrt{v_Y(t, T, T + \delta)}X_Y} e^{-\frac{1}{2}X_Y^2} dX_Y. \end{aligned}$$

Now, it is noted that

$$-\frac{1}{2}(X_Y - \sqrt{v_Y(t, T, T + \delta)})^2 + \frac{1}{2}v_Y(t, T, T + \delta) = -\frac{1}{2}X_Y^2 + \sqrt{v_Y(t, T, T + \delta)}X_Y,$$

which implies that we can integrate by substitution with $x_Y = X_Y - \sqrt{v_Y(t, T, T + \delta)}$ whereby $dx_Y = dX_Y$ and the intervals to be integrated over change to

$$\begin{aligned} x_Y^{top} &= X_Y^{top} - \sqrt{v_Y(t, T, T + \delta)} = \infty, \\ x_Y^{bottom} &= X_Y^{bottom} - \sqrt{v_Y(t, T, T + \delta)} = \frac{\ln K - m_Y(t, T, T + \delta) - v_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}. \end{aligned}$$

Thus, the first payment expectation can be calculated as

$$\begin{aligned} C_1(t, T, T + \delta; K) &= P(t, T) \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln K - m_Y(t, T, T + \delta) - v_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}}^{\infty} e^{m_Y(t, T, T + \delta) + \frac{1}{2}v_Y(t, T, T + \delta)} e^{-\frac{1}{2}x_Y^2} dx_Y \\ &= P(t, T) e^{m_Y(t, T, T + \delta) + \frac{1}{2}v_Y(t, T, T + \delta)} \Phi \left(\frac{m_Y(t, T, T + \delta) + v_Y(t, T, T + \delta) - \ln K}{\sqrt{v_Y(t, T, T + \delta)}} \right). \end{aligned}$$

Due to the property of the log-normal distribution, it follows that

$$E_t^{Q^T} [e^{Y(T, T, T + \delta)}] = E_t^{Q^T} [Z(T, T, T + \delta)] = e^{m_Y(t, T, T + \delta) + \frac{1}{2}v_Y(t, T, T + \delta)}.$$

Since $Z(t, T, T + \delta)$ is a Q^T -martingale, this implies that

$$Z(t, T, T + \delta) = \frac{P(t, T + \delta)}{P(t, T)} = e^{m_Y(t, T, T + \delta) + \frac{1}{2}v_Y(t, T, T + \delta)}.$$

Now, insert that in the expression above to obtain

$$C_1(t, T, T + \delta; K) = P(t, T + \delta) \Phi \left(\frac{m_Y(t, T, T + \delta) + v_Y(t, T, T + \delta) - \ln K}{\sqrt{v_Y(t, T, T + \delta)}} \right).$$

To summarize, the call option with maturity at T and strike price K written on the zero-coupon bond maturing at $T + \delta$ is given by

$$\begin{aligned} C(t, T, T + \delta; K) &= C_1(t, T, T + \delta; K) + C_2(t, T, T + \delta; K) \\ &= P(t, T + \delta) \Phi(d_1) - KP(t, T) \Phi(d_2), \end{aligned}$$

where

- $d_1 = \frac{m_Y(t, T, T + \delta) + v_Y(t, T, T + \delta) - \ln K}{\sqrt{v_Y(t, T, T + \delta)}}$,
- $d_2 = d_1 - \sqrt{v_Y(t, T, T + \delta)}$.

The conditional mean of $Y(T, T, T + \delta)$ under the T -forward measure is

$$m_Y(t, T, T + \delta) = Y(t, T, T + \delta) - \frac{1}{2} \int_t^T \sum_{j=1}^3 (\Sigma' [B(s, T + \delta) - B(s, T)] [B(s, T + \delta) - B(s, T)]' \Sigma)_{j,j} ds,$$

while its conditional variance is given by

$$\begin{aligned} v_Y(t, T, T + \delta) &= V^{QT} \left[\int_t^T [B(s, T + \delta) - B(s, T)]' \Sigma dW_s^{QT} | \mathcal{F}_t \right] \\ &= \int_t^T \sum_{j=1}^3 (\Sigma' [B(s, T + \delta) - B(s, T)] [B(s, T + \delta) - B(s, T)]' \Sigma)_{j,j} ds. \end{aligned}$$

From this it follows that

$$\begin{aligned} m_Y(t, T, T + \delta) &= Y(t, T, T + \delta) - \frac{1}{2} v_Y(t, T, T + \delta) \\ &= \ln \left(\frac{P(t, T + \delta)}{P(t, T)} \right) - \frac{1}{2} v_Y(t, T, T + \delta). \end{aligned}$$

This implies that we can rewrite d_1 as

$$d_1 = \frac{\ln \left(\frac{P(t, T + \delta)}{P(t, T)K} \right) + \frac{1}{2} v_Y(t, T, T + \delta)}{\sqrt{v_Y(t, T, T + \delta)}}.$$

This structure is consistent with the more simple Gaussian option price formulas derived in Jamshidian (1989) and Chen (1992). Importantly, once we have the analytical formula for the conditional variance, we have all ingredients needed to calculate the call option price. This is the task we now turn to.

The Analytical Formula for the Conditional Variance

To begin, we expand the expression for the conditional variance of $Y(t, T, T + \delta)$ as follows:

$$\begin{aligned}
v_Y(t, T, T + \delta) &= \int_t^T \sum_{j=1}^3 \left[\begin{pmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ 0 & \sigma_{22} & \sigma_{32} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} B^1(s, T + \delta) - B^1(s, T) \\ B^2(s, T + \delta) - B^2(s, T) \\ B^3(s, T + \delta) - B^3(s, T) \end{pmatrix} \right. \\
&\quad \times \begin{pmatrix} B^1(s, T + \delta) - B^1(s, T) & B^2(s, T + \delta) - B^2(s, T) & B^3(s, T + \delta) - B^3(s, T) \end{pmatrix} \\
&\quad \left. \times \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \right]_{j,j} ds.
\end{aligned}$$

This produces a total of six unique integrals that have to be calculated.

The first of the six integrals is given by

$$v_Y^1(t, T, T + \delta) = \sigma_{11}^2 \int_t^T [B^1(s, T + \delta) - B^1(s, T)]^2 ds.$$

The second integral is given by

$$v_Y^2(t, T, T + \delta) = (\sigma_{21}^2 + \sigma_{22}^2) \int_t^T [B^2(s, T + \delta) - B^2(s, T)]^2 ds.$$

The third integral is given by

$$v_Y^3(t, T, T + \delta) = (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \int_t^T [B^3(s, T + \delta) - B^3(s, T)]^2 ds.$$

The fourth integral is given by

$$v_Y^4(t, T, T + \delta) = 2\sigma_{11}\sigma_{21} \int_t^T [B^1(s, T + \delta) - B^1(s, T)][B^2(s, T + \delta) - B^2(s, T)] ds.$$

The fifth integral is given by

$$v_Y^5(t, T, T + \delta) = 2\sigma_{11}\sigma_{31} \int_t^T [B^1(s, T + \delta) - B^1(s, T)][B^3(s, T + \delta) - B^3(s, T)] ds.$$

The sixth and final integral is given by

$$v_Y^6(t, T, T + \delta) = 2(\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \int_t^T [B^2(s, T + \delta) - B^2(s, T)][B^3(s, T + \delta) - B^3(s, T)] ds.$$

Unreported calculations show that the conditional volatility of the $Y(t, T, T + \delta)$ process is¹⁸

$$\begin{aligned}
v_Y(t, T, T + \delta) &= \sum_{i=1}^6 v_Y^i(t, T, T + \delta) \\
&= \sigma_{11}^2 \delta^2 (T - t) + (\sigma_{21}^2 + \sigma_{22}^2) \left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) e^{-2\lambda \delta} \left[\frac{\delta^2 - (T + \delta - t)^2 e^{-2\lambda(T-t)}}{2\lambda} + \frac{\delta - (T + \delta - t) e^{-2\lambda(T-t)}}{2\lambda^2} + \frac{1 - e^{-2\lambda(T-t)}}{4\lambda^3} \right] \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \left[-\frac{1}{2\lambda} (T - t)^2 e^{-2\lambda(T-t)} - \frac{1}{2\lambda^2} (T - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{4\lambda^3} \right] \\
&\quad - (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \frac{(1 - e^{-\lambda \delta}) e^{-\lambda \delta}}{\lambda^2} \left[\delta - (T + \delta - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \frac{1 - e^{-\lambda \delta}}{\lambda^2} \left[\frac{1 - e^{-2\lambda(T-t)}}{2\lambda} - (T - t) e^{-2\lambda(T-t)} \right] \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \frac{1}{\lambda} \delta e^{-\lambda \delta} \left[(T - t) e^{-2\lambda(T-t)} - \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad + (\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2) \frac{1}{\lambda} e^{-\lambda \delta} \left[(T - t)^2 e^{-2\lambda(T-t)} + \frac{1}{\lambda} (T - t) e^{-2\lambda(T-t)} - \frac{1 - e^{-2\lambda(T-t)}}{2\lambda^2} \right] \\
&\quad + 2\sigma_{11}\sigma_{21}\delta(1 - e^{-\lambda \delta}) \frac{1 - e^{-\lambda(T-t)}}{\lambda^2} \\
&\quad + 2\sigma_{11}\sigma_{31}\delta \left[-\frac{1}{\lambda} (T - t) e^{-\lambda(T-t)} - \frac{1}{\lambda} e^{-\lambda \delta} (\delta - (T + \delta - t) e^{-\lambda(T-t)}) + 2(1 - e^{-\lambda \delta}) \frac{1 - e^{-\lambda(T-t)}}{\lambda^2} \right] \\
&\quad + (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \left(\frac{1 - e^{-\lambda \delta}}{\lambda} \right)^2 \frac{1 - e^{-2\lambda(T-t)}}{\lambda} \\
&\quad + (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \frac{1}{\lambda^2} e^{-2\lambda \delta} \left[\delta - (T + \delta - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad + (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \frac{1}{\lambda^2} \left[- (T - t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{2\lambda} \right] \\
&\quad - (\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}) \frac{1}{\lambda^2} e^{-\lambda \delta} \left[\delta - (2T + \delta - 2t) e^{-2\lambda(T-t)} + \frac{1 - e^{-2\lambda(T-t)}}{\lambda} \right].
\end{aligned}$$

Appendix B: Kalman Filter Estimation of Shadow-Rate Models

In this appendix we describe the estimation of the shadow-rate models based on the extended Kalman filter.

For affine Gaussian models, in general, the conditional mean vector and the conditional covariance matrix are

$$\begin{aligned}
E^P[X_T | \mathcal{F}_t] &= (I - \exp(-K^P \Delta t)) \theta^P + \exp(-K^P \Delta t) X_t, \\
V^P[X_T | \mathcal{F}_t] &= \int_0^{\Delta t} e^{-K^P s} \Sigma \Sigma' e^{-(K^P)' s} ds,
\end{aligned}$$

where $\Delta t = T - t$. We compute conditional moments of discrete observations and obtain the state transition equation

$$X_t = (I - \exp(-K^P \Delta t)) \theta^P + \exp(-K^P \Delta t) X_{t-1} + \xi_t,$$

where Δt is the time between observations. In the standard Kalman filter, the measurement equation would

¹⁸The calculations leading to this result are available from the authors upon request.

be affine, in which case

$$y_t = A + BX_t + \varepsilon_t.$$

The assumed error structure is

$$\begin{pmatrix} \xi_t \\ \varepsilon_t \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right],$$

where the matrix H is assumed diagonal, while the matrix Q has the following structure:

$$Q = \int_0^{\Delta t} e^{-K^P s} \Sigma \Sigma' e^{-(K^P)' s} ds.$$

In addition, the transition and measurement errors are assumed orthogonal to the initial state.

Now we consider Kalman filtering, which we use to evaluate the likelihood function.

Due to the assumed stationarity, the filter is initialized at the unconditional mean and variance of the state variables under the P -measure: $X_0 = \theta^P$ and $\Sigma_0 = \int_0^\infty e^{-K^P s} \Sigma \Sigma' e^{-(K^P)' s} ds$, which we calculate using the analytical solutions provided in Fisher and Gilles (1996).

Denote the information available at time t by $Y_t = (y_1, y_2, \dots, y_t)$, and denote model parameters by ψ . Consider period $t - 1$ and suppose that the state update X_{t-1} and its mean square error matrix Σ_{t-1} have been obtained. The prediction step is

$$X_{t|t-1} = E^P[X_t|Y_{t-1}] = \Phi_t^{X,0}(\psi) + \Phi_t^{X,1}(\psi)X_{t-1},$$

$$\Sigma_{t|t-1} = \Phi_t^{X,1}(\psi)\Sigma_{t-1}\Phi_t^{X,1}(\psi)' + Q_t(\psi),$$

where $\Phi_t^{X,0} = (I - \exp(-K^P \Delta t))\theta^P$, $\Phi_t^{X,1} = \exp(-K^P \Delta t)$, and $Q_t = \int_0^{\Delta t} e^{-K^P s} \Sigma \Sigma' e^{-(K^P)' s} ds$, while Δt is the time between observations.

In the time- t update step, $X_{t|t-1}$ is improved by using the additional information contained in Y_t . We have

$$X_t = E[X_t|Y_t] = X_{t|t-1} + \Sigma_{t|t-1}B(\psi)'F_t^{-1}v_t,$$

$$\Sigma_t = \Sigma_{t|t-1} - \Sigma_{t|t-1}B(\psi)'F_t^{-1}B(\psi)\Sigma_{t|t-1},$$

where

$$v_t = y_t - E[y_t|Y_{t-1}] = y_t - A(\psi) - B(\psi)X_{t|t-1},$$

$$F_t = \text{cov}(v_t) = B(\psi)\Sigma_{t|t-1}B(\psi)' + H(\psi),$$

$$H(\psi) = \text{diag}(\sigma_\varepsilon^2(\tau_1), \dots, \sigma_\varepsilon^2(\tau_N)).$$

At this point, the Kalman filter has delivered all ingredients needed to evaluate the Gaussian log likelihood, the prediction-error decomposition of which is

$$\log l(y_1, \dots, y_T; \psi) = \sum_{t=1}^T \left(-\frac{N}{2} \log(2\pi) - \frac{1}{2} \log |F_t| - \frac{1}{2} v_t' F_t^{-1} v_t \right),$$

where N is the number of observed yields. We numerically maximize the likelihood with respect to ψ using the Nelder-Mead simplex algorithm. Upon convergence, we obtain standard errors from the estimated covariance matrix,

$$\widehat{\Omega}(\widehat{\psi}) = \frac{1}{T} \left[\frac{1}{T} \sum_{t=1}^T \frac{\partial \log l_t(\widehat{\psi})}{\partial \psi} \frac{\partial \log l_t(\widehat{\psi})'}{\partial \psi} \right]^{-1},$$

where $\widehat{\psi}$ denotes the estimated model parameters.

This completes the description of the standard Kalman filter. However, in the shadow-rate models, the zero-coupon bond yields are not affine functions of the state variables. Instead, the measurement equation takes the general form

$$y_t = z(X_t; \psi) + \varepsilon_t.$$

In the extended Kalman filter we use, this equation is linearized through a first-order Taylor expansion around the best guess of X_t in the prediction step of the Kalman filter algorithm. Thus, in the notation introduced above, this best guess is denoted $X_{t|t-1}$ and the approximation is given by

$$z(X_t; \psi) \approx z(X_{t|t-1}; \psi) + \left. \frac{\partial z(X_t; \psi)}{\partial X_t} \right|_{X_t=X_{t|t-1}} (X_t - X_{t|t-1}).$$

Now, by defining

$$A_t(\psi) \equiv z(X_{t|t-1}; \psi) - \left. \frac{\partial z(X_t; \psi)}{\partial X_t} \right|_{X_t=X_{t|t-1}} X_{t|t-1} \quad \text{and} \quad B_t(\psi) \equiv \left. \frac{\partial z(X_t; \psi)}{\partial X_t} \right|_{X_t=X_{t|t-1}},$$

the measurement equation can be given in an affine form as

$$y_t = A_t(\psi) + B_t(\psi)X_t + \varepsilon_t,$$

and the steps in the algorithm proceeds as previously described.

References

- Bauer, Michael D., Glenn D. Rudebusch, and Jing (Cynthia) Wu, 2012, “Correcting Estimation Bias in Dynamic Term Structure Models,” *Journal of Business and Economic Statistics*, Vol. 30, No. 3, 454-467.
- Black, Fisher, 1995, “Interest Rates as Options,” *Journal of Finance*, Vol. 50, No. 7, 1371-1376.
- Bomfim, Antulio N., 2003, “‘Interest Rates as Options:’ Assessing the markets’ view of the liquidity trap,” Working Paper 2003-45, Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.
- Chen, Ren-Raw, 1992, “Exact Solutions for Futures and European Futures Options on Pure Discount Bonds,” *Journal of Financial and Quantitative Analysis*, Vol. 27, No. 1, 97-107.
- Cheridito, Patrick, Damir Filipović, and Robert L. Kimmel, 2007, “Market Price of Risk Specifications for Affine Models: Theory and Evidence,” *Journal of Financial Economics*, Vol. 83, 123-170.
- 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*, Vol. 164, 4-20.
- Christensen, Jens H. E. and James M. Gillan, 2012, “Could the U.S. Treasury Benefit from Issuing More TIPS?,” Working Paper 2011-16, Federal Reserve Bank of San Francisco.
- Christensen, Jens H. E., Jose A. Lopez, and Glenn D. Rudebusch, 2010, “Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields,” *Journal of Money, Credit and Banking*, Supplement to Vol. 42, no. 6, 143-178.
- Christensen, Jens H. E. and Glenn D. Rudebusch, 2012, “The Response of Interest Rates to U.S. and U.K. Quantitative Easing,” *Economic Journal*, Vol. 122, 385-414.
- Diebold, Francis X. and Glenn D. Rudebusch, 2013, *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*, Princeton, NJ: Princeton University Press.
- Duffee, Gregory R., 2002, “Term Premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, Vol. 57, 405-443.
- Duffie, Darrell and Rui Kan, 1996, “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, Vol. 6, 379-406.

- Fisher, Mark and Christian Gilles, 1996, "Term Premia in Exponential-Affine Models of the Term Structure," Manuscript, Board of Governors of the Federal Reserve System.
- Gorovoi, Viatcheslav and Vadim Linetsky, 2004, "Black's Model of Interest Rates as Options, Eigenfunction Expansions and Japanese Interest Rates," *Mathematical Finance*, Vol. 14, No. 1, 49-78.
- Ichiue, Hibiki and Yoichi Ueno, 2007, "Equilibrium Interest Rates and the Yield Curve in a Low Interest Rate Environment," Working Paper 2007-E-18, Bank of Japan.
- Jamshidian, Farshid, 1989, "An Exact Bond Option Formula," *Journal of Finance*, Vol. 44, No. 1, 205-209.
- Kim, Don H. and Kenneth J. Singleton, 2012, "Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields," *Journal of Econometrics*, Vol. 170, 32-49.
- Krippner, Leo, 2012, "Modifying Gaussian Term Structure Models when Interest Rates Are Near the Zero Lower Bound," Discussion Paper 2012-02, Reserve Bank of New Zealand.
- Krippner, Leo, 2013, "Measuring the Stance of Monetary Policy in Zero Lower Bound Environments," *Economics Letters*, Vol. 118, 135-138.
- Litterman, R. and J. A. Scheinkman, 1991, "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, Vol. 1, 62-74.
- Merton, Robert C., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, Vol. 29, No. 2, 449-470.
- Nelson, Charles R. and Andrew F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Business*, Vol. 60, 473-489.
- Ueno, Yoichi, Naohiko Baba, and Yuji Sakurai, 2006, "The Use of the Black Model of Interest Rates as Options for Monitoring the JGB Market Expectations," Working Paper 2006-E-15, Bank of Japan.
- Vasiček, Oldrich, 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, Vol. 5, 177-188.