Swiss Finance Institute
Research Paper Series N°14-15

Linear-Rational Term Structure Models

Damir FILIPOVIC
Ecole Polytechnique Fédérale de Lausanne and Swiss Finance Institute

Martin LARSSON
Ecole Polytechnique Fédérale de Lausanne and Swiss Finance Institute

Anders TROLLE
Ecole Polytechnique Fédérale de Lausanne and Swiss Finance Institute
Linear-Rational Term Structure Models *

Damir Filipović Martin Larsson Anders Trolle
EPFL and Swiss Finance Institute
21 February, 2014

Abstract

We introduce the class of linear-rational term structure models, where the state price density is modeled such that bond prices become linear-rational functions of the current state. This class is highly tractable with several distinct advantages: i) ensures nonnegative interest rates, ii) easily accommodates unspanned factors affecting volatility and risk premia, and iii) admits analytical solutions to swaptions. For comparison, exponential-affine term structure models can match either i) or ii), but not both simultaneously, and never iii). A parsimonious specification of the model with three term structure factors and at least two unspanned factors has a very good fit to both interest rate swaps and swaptions since 1997. In particular, the model captures well the dynamics of risk premia as well as the dynamics of the term structure and volatility during the recent period of near-zero interest rates.

Keywords: Swaps, Swaptions, Unspanned Factors, Zero Lower Bound

JEL Classification: E43, G12, G13

*The authors wish to thank seminar participants at the 10th German Probability and Statistics Days in Mainz, the Conference on Stochastic Analysis and Applications in Lausanne, the Seventh Bachelier Colloquium on Mathematical Finance and Stochastic Calculus in Metabief, Conference of Current Topics in Mathematical Finance in Vienna, the Princeton-Lausanne Workshop on Quantitative Finance in Princeton, the 29:th European Meeting of Statisticians in Budapest, the Term Structure Modeling at the Zero Lower Bound workshop at the Federal Reserve Bank of San Francisco, the Cambridge Finance Seminar, the London Mathematical Finance Seminar, the Seminar on Mathematical Finance in Vienna, and Ken Singleton (discussant) and Pierre Collin-Dufresne for their comments. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 307465-POLYTE.
1 Introduction

The current environment with near-zero interest rates creates difficulties for many existing term structure models, most notably Gaussian models that invariably place large probabilities on negative future rates. Models that respect the zero lower bound on interest rates exist but are often restricted in their ability to accommodate unspanned factors affecting volatility and risk premia and to price many types of interest rate derivatives. In light of these limitations, the purpose of this paper is twofold: First, we introduce a new class of term structure models, the linear-rational, which is highly tractable and i) ensures nonnegative interest rates, ii) easily accommodates unspanned factors affecting volatility and risk premia, and iii) admits analytical solutions to swaptions—an important class of interest rate derivatives that underlie the pricing and hedging of mortgage-backed securities, callable agency securities, life insurance products, and a wide variety of structured products. Second, we perform an extensive empirical analysis, focusing in particular on the recent period of near-zero interest rates.

The first contribution of the paper is to introduce the class of linear-rational term structure models. A sufficient condition for the absence of arbitrage opportunities in a model of a financial market is the existence of a state price density: a positive adapted process \( \zeta_t \) such that the price \( \Pi(t, T) \) at time \( t \) of any time \( T \) cash-flow, \( C_T \) say, is given by

\[
\Pi(t, T) = \frac{1}{\zeta_t} \mathbb{E}[\zeta_T C_T \mid \mathcal{F}_t],
\]

where we suppose there is a filtered probability space \( (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}) \) on which all random quantities are defined. Following Constantinides (1992), our approach to modeling the term structure is to directly specify the state price density. Specifically, we assume a multivariate factor process with a linear drift, and a state price density, which is a linear function of the current state. In this case, zero-coupon bond prices and the short rate become linear-rational functions of the current state, which is why we refer to the framework as linear-rational. One attractive feature of the framework is that one can easily ensure nonnegative interest rates. Another attractive feature is that the martingale component of the factor process does not affect the term structure. This implies that one can easily allow for factors that affect prices of interest rate derivatives without affecting bond prices. Assuming that the factor process has diffusive dynamics, we show that the state vector can be partitioned into factors that affect the term structure, factors that affect interest rate volatility but not the term structure (unspanned stochastic volatility, or USV, factors), and factors that neither affect the term structure nor interest rate volatility but may nevertheless
have an indirect impact on interest rate derivatives. Assuming further that the factor process is of the square-root type, we show how swaptions can be priced analytically. This specific model is termed the linear-rational square-root (LRSQ) model.

The second contribution of the paper is an extensive empirical analysis of the LRSQ model. We utilize a panel data set consisting of term structures of swap rates and swaption implied volatilities. The sample period is from January 1997 to August 2013, and the estimation approach is quasi-maximum likelihood in conjunction with the Kalman filter. A specification of the model with three term structure factors and at least two USV factors has a very good fit to both interest rate swaps and swaptions simultaneously. This holds true also for the part of the sample period where short-term rates were very close to the zero lower bound.

We investigate the inherent properties of the model using long samples of simulated data. The model captures several important characteristics of risk premia in the swap market. In line with the data, the unconditional mean and volatility of excess returns increase with swap maturity, but in such a way that unconditional Sharpe ratios decrease with swap maturity. Furthermore, in contrast to many existing studies, in our swap data volatility predicts excess returns, while the predictive power of the slope of the term structure is much weaker. The model largely captures these results.

The model also replicates important features of term structure dynamics at the zero lower bound. Consistent with the data, the model generates extended periods of near-zero short rates as well as highly asymmetric distributions of future short rates, with the most likely value of future short rates being significantly lower than the mean value. In addition, the model captures how the main term structure factor changes from being a “level” factor during normal times to being more of a “slope” factor during times of near-zero short rates.

By design the model exhibits USV, but also quantitatively the model is able to match the degree of USV observed in the data. More importantly, the model captures the empirical regularity that volatility becomes gradually more level-dependent as the underlying interest rate approaches the zero lower bound.

A special case of our general linear-rational framework is the model considered by Carr, Gabaix, and Wu (2009). However, the factor process in their model is time-inhomogeneous and non-stationary, while the LRSQ model that we evaluate empirically is time-homogeneous and stationary. Furthermore, the volatility structure in their model is very different from the one in the LRSQ model.

1The historical mean excess returns and Sharpe ratios are inflated by the downward trend in interest rates over the sample period and, indeed, the model-implied values are lower.

2More generally, the linear-rational framework is related to the frameworks in Rogers (1997) and
The exponential-affine framework, see, e.g., Duffie and Kan (1996) and Dai and Singleton (2000), is arguably the dominant one in the term structure literature. In this framework, one can either ensure nonnegative interest rates (which requires all factors to be of the square-root type) or accommodate USV (which requires at least one conditionally Gaussian factor), but not both. Furthermore, no exponential-affine model admits analytical solutions to swaptions. In contrast, the linear-rational framework accommodates all three features.

The paper is structured as follows. Section 2 lays out the general framework, leaving the martingale term of the factor process unspecified. Section 3 specializes to the case where the factor process has diffusive dynamics. Section 4 further specializes to the case where the factor process is of the square-root type. Section 5 discusses a flexible specification of market prices of risk. Section 6 describes the data. Section 7 presents the empirical results. Section 8 concludes. All proofs are given in the appendix.

2 The Linear-Rational Framework

In this section the linear-rational framework is introduced, and explicit formulas for zero-coupon bond prices and short rate are presented. We then discuss how unspanned factors arise in this setting, and how the factor process after a change of coordinates can be decomposed into spanned and unspanned components. We then describe interest rate swaptions, and derive a swaption pricing formula. Finally, the linear-rational framework is compared and contrasted with existing models.

2.1 Term Structure Specification

A linear-rational term structure model consists of two components: a multivariate factor process \( X_t \) whose state space is some subset \( E \subset \mathbb{R}^d \), and a state price density \( \zeta_t \) given as a deterministic function of the current state. The linear-rational class becomes tractable due to the interplay between two basic structural assumptions we impose on these components: the factor process has a linear drift, and the state price density is a linear function of the current state. More specifically, we assume that \( X_t \) is of the form

\[
dX_t = \kappa(\theta - X_t)dt + dM_t
\]

Flesaker and Hughston (1996).

Alternatively, the “shadow rate” model of Black (1995) ensures nonnegative interest rates; see, e.g., Kim and Singleton (2012), Bauer and Rudebusch (2013), and Christensen and Rudebusch (2013) for recent applications of this framework.
for some $\kappa \in \mathbb{R}^{d \times d}$, $\theta \in \mathbb{R}^d$, and some martingale $M_t$. Typically $X_t$ will follow Markovian dynamics, although this is not necessary for this section. Next, the state price density is assumed to be given by

$$\zeta_t = e^{-\alpha t} (\phi + \psi^T X_t),$$

for some $\phi \in \mathbb{R}$ and $\psi \in \mathbb{R}^d$ such that $\phi + \psi^T x > 0$ for all $x \in E$, and some $\alpha \in \mathbb{R}$. As we discuss below, the role of the parameter $\alpha$ is to ensure that the short rate stays nonnegative.

The linear drift of the factor process implies that conditional expectations have the following simple form, as can be seen from Lemma A.1:

$$\mathbb{E}[X_T \mid \mathcal{F}_t] = \theta + e^{-\kappa(T-t)}(X_t - \theta), \quad t \leq T.$$  \hfill (4)

An immediate consequence is that the zero-coupon bond prices and the short rate become linear-rational functions of the current state, which is why we refer to this framework as linear-rational. Indeed, the basic pricing formula (1) with $C = 1$ shows that the zero-coupon bond prices are given by $P(t,T) = F(T-t, X_t)$, where

$$F(\tau, x) = \frac{(\phi + \psi^T \theta)e^{-\alpha \tau} + \psi^T e^{-(\alpha + \kappa)\tau}(x - \theta)}{\phi + \psi^T x}. \hfill (5)$$

The short rate is then obtained via the formula $r_t = -\partial_T \log P(t,T)|_{T=t}$, and is given by

$$r_t = \alpha - \frac{\psi^T \kappa(\theta - X_t)}{\phi + \psi^T X_t}. \hfill (6)$$

The latter expression clarifies the role of the parameter $\alpha$; provided that the short rate is bounded from below, we may guarantee that it stays nonnegative by choosing $\alpha$ large enough. This leads to an intrinsic choice of $\alpha$ as the smallest value that yields a nonnegative short rate. In other words, we define

$$\alpha^* = \sup_{x \in E} \frac{\psi^T \kappa(\theta - x)}{\phi + \psi^T x} \quad \text{and} \quad \alpha_* = \inf_{x \in E} \frac{\psi^T \kappa(\theta - x)}{\phi + \psi^T x}, \hfill (7)$$

and set $\alpha = \alpha^*$, provided this is finite. The short rate then satisfies

$$r_t \in [0, \alpha^* - \alpha_*] \quad (r_t \in [0, \infty) \text{ if } \alpha_* = -\infty).$$

4One could replace the drift $\kappa(\theta - X_s)$ in (2) with the slightly more general form $b + \beta X_s$ for some $b \in \mathbb{R}^d$ and $\beta \in \mathbb{R}^{d \times d}$. The gain in generality is moderate (the two parameterizations are equivalent if $b$ lies in the range of $\beta$) and is trumped by the gain in notational clarity that will be achieved by using the form (2). The latter form also has the advantage of allowing for a “mean-reversion” interpretation of the drift.
Notice that \( \alpha^* \) and \( \alpha_* \) depend on the parameters of the process \( X_t \), which are determined through calibration. A crucial step of the model validation process is therefore to verify that the range of possible short rates is sufficiently wide. Finally, notice that whenever the eigenvalues of \( \kappa \) have nonnegative real part, one easily verifies the equality
\[
\lim_{\tau \to \infty} -\frac{1}{\tau} \log F(\tau, x) = \alpha,
\]
valid for any \( x \in E \). In other words, \( \alpha \) can be interpreted as an infinite-maturity forward rate.

### 2.2 Unspanned Factors

Our focus is now to describe the directions \( \xi \in \mathbb{R}^d \) such that the term structure remains unchanged when the state vector moves along \( \xi \). It is convenient to carry out this discussion in terms of the kernel of a function.\(^5\)

**Definition 2.1.** The term structure kernel, denoted by \( \mathcal{U} \), is given by
\[
\mathcal{U} = \bigcap_{\tau \geq 0} \ker F(\tau, \cdot).
\]

That is, \( \mathcal{U} \) consists of all \( \xi \in \mathbb{R}^d \) such that \( \nabla F(\tau, x)^	op \xi = 0 \) for all \( \tau \geq 0 \) and all \( x \in E \).\(^6\) Therefore the location of the state \( X_t \) along the direction \( \xi \) cannot be recovered solely from knowledge of the time \( t \) bond prices \( P(t, t + \tau), \tau \geq 0 \). In this sense the term structure kernel is unspanned by the term structure. In Section 3.1 we will discuss how this notion relates to spanning in the sense of bond market completeness. The following result characterizes \( \mathcal{U} \) in terms of the model parameters.

**Theorem 2.2.** Assume the term structure is not trivial.\(^7\) Then \( \mathcal{U} \) is the largest

\(^5\)We define the kernel of a differentiable function \( f \) on \( E \) by
\[
\ker f = \{ \xi \in \mathbb{R}^d : \nabla f(x)^	op \xi = 0 \text{ for all } x \in E \}.
\]
This notion generalizes the standard one: if \( f(x) = v^	op x \) is linear, for some \( v \in \mathbb{R}^d \), then \( \nabla f(x) = v \) for all \( x \in E \), so \( \ker f = \ker v^	op \) coincides with the usual notion of kernel.

\(^6\)Here and in the sequel, \( \nabla F(\tau, x) \) denotes the gradient with respect to the \( x \) variables.

\(^7\)We say that the term structure is trivial if the short rate \( r_t \) is constant. In view of (6), this happens if and only if \( \psi \) is an eigenvector of \( \kappa^	op \) with eigenvalue \( \lambda \) satisfying \( \lambda(\phi + \psi^	op \theta) = 0 \). In this case, we have \( r_t \equiv \alpha + \lambda \) and \( \mathcal{U} = \mathbb{R}^d \), while the right side of (8) equals \( \ker \psi^	op \). The assumption that the term structure be not trivial will be in force throughout the paper.
subspace of \( \ker \psi^\top \) that is invariant under \( \kappa \). Formally, this is equivalent to

\[
\mathcal{U} = \bigcap_{p=0}^{d-1} \ker \psi^\top \kappa^p.
\]  

(8)

In the case where \( \kappa \) is diagonalizable, this leads to the following corollary.

**Corollary 2.3.** Assume \( \kappa \) is diagonalizable with real eigenvalues, i.e. \( \kappa = S^{-1} \Lambda S \) with \( S \) invertible and \( \Lambda \) diagonal and real. Then \( \mathcal{U} = \{0\} \) if and only if all eigenvalues of \( \kappa \) are distinct and all components of \( S^{-\top} \psi \) are nonzero.

We now transform the state space so that the unspanned directions correspond to the last components of the state vector. To this end, first let \( S \) be any invertible linear transformation on \( \mathbb{R}^d \). The transformed factor process \( \hat{X}_t = SX_t \) satisfies the linear drift dynamics

\[
d\hat{X}_t = \hat{\kappa}(\hat{\theta} - \hat{X}_t)dt + d\hat{M}_t,
\]

where

\[
\hat{\kappa} = S\kappa S^{-1}, \quad \hat{\theta} = S\theta, \quad \hat{M}_t = SM_t.
\]  

(9)

Defining also

\[
\hat{\phi} = \phi, \quad \hat{\psi} = S^{-\top} \psi,
\]  

(10)

we have \( \zeta_t = e^{-\alpha t}(\hat{\phi} + \hat{\psi}^\top \hat{X}_t) \). This gives a linear-rational term structure model that is equivalent to the original one. Suppose now that \( S \) maps the term structure kernel into the standard basis of \( \mathbb{R}^d = \mathbb{R}^m \times \mathbb{R}^n \),

\[
S(\mathcal{U}) = \{0\} \times \mathbb{R}^n
\]  

(11)

where \( n = \dim \mathcal{U} \) and \( m = d - n \). Decomposing the transformed factor process accordingly, \( \hat{X}_t = (Z_t, U_t) \), our next result and the subsequent discussion will show that \( Z_t \) affects the term structure, while \( U_t \) does not.

**Theorem 2.4.** Let \( m, n \geq 0 \) be integers with \( m + n = d \). Then (11) holds if and only if the transformed model parameters (9)–(10) satisfy:

\[
\text{(i) } \hat{\psi} = (\hat{\psi}_Z, 0) \in \mathbb{R}^m \times \mathbb{R}^n; \]

\[
\text{(ii) } \hat{\phi} = \phi, \quad \hat{\psi} = S^{-\top} \psi,
\]

\[
\text{(iii) } \ker \hat{\psi}^\top \text{ is a subspace of } \ker \psi^\top \text{ that is invariant under } \hat{\kappa}. \quad \text{Condition (iii) asserts that } \{0\} \times \mathbb{R}^n \text{ is the largest subspace of } \ker \psi^\top \text{ with this property.}
\]

\[7\]
(ii) \( \hat{\kappa} \) has block lower triangular structure, 
\[
\hat{\kappa} = \begin{pmatrix}
\hat{\kappa}_{ZZ} & 0 \\
\hat{\kappa}_{UZ} & \hat{\kappa}_{UU}
\end{pmatrix} \in \mathbb{R}^{(m+n) \times (m+n)};
\]

(iii) The upper left block \( \hat{\kappa}_{ZZ} \) of \( \hat{\kappa} \) satisfies
\[
\bigcap_{p=0}^{m-1} \ker \hat{\psi}_Z^{\top} \hat{\kappa}_{ZZ}^p = \{0\}.
\]

Assuming that (11) holds, and writing \( Sx = (z, u) \in \mathbb{R}^m \times \mathbb{R}^n \) and \( \hat{\theta} = (\hat{\theta}_Z, \hat{\theta}_U) \), we now see that
\[
\hat{F}(\tau, z) = F(\tau, x) = \frac{(\hat{\phi} + \hat{\psi}_Z^{\top} \hat{\theta}_Z)e^{-\alpha \tau} + \hat{\psi}_Z^{\top} e^{- (\alpha + \hat{\kappa}_{ZZ}) \tau} (z - \hat{\theta}_Z)}{\hat{\phi} + \hat{\psi}_Z^{\top} z}
\]
does not depend on \( u \). Hence the bond prices are given by \( P(t, T) = \hat{F}(T - t, Z_t) \). This gives a clear interpretation of the components of \( U_t \) as unspanned factors: their values do not influence the current term structure. As a consequence, a snapshot of the term structure at time \( t \) does not provide any information about \( U_t \). The sub-vector \( Z_t \), on the other hand, directly impacts the term structure, and can be reconstructed from a snapshot of the term structure at time \( t \), under mild technical conditions. For this reason we refer to the components of \( Z_t \) as term structure factors. The following theorem formalizes the above discussion.

**Theorem 2.5.** The term structure \( \hat{F}(\tau, z) \) is injective if and only if \( \hat{\kappa}_{ZZ} \) is invertible and \( \hat{\phi} + \hat{\psi}_Z^{\top} \hat{\theta}_Z \neq 0 \).

In view of Theorem 2.4, the dynamics of \( \hat{X}_t = (Z_t, U_t) \) can be decomposed into term structure dynamics
\[
dZ_t = \hat{\kappa}_{ZZ} (\hat{\theta}_Z - Z_t) dt + d\hat{M}_{Zt} \tag{12}
\]
and unspanned factor dynamics
\[
dU_t = \left( \hat{\kappa}_{UZ}(\hat{\theta}_Z - Z_t) + \hat{\kappa}_{UU}(\hat{\theta}_U - U_t) \right) dt + d\hat{M}_{U_t}
\]

---

Footnote 9: Injectivity means that if \( \hat{F}(\tau, z) = \hat{F}(\tau, z') \) for all \( \tau \geq 0 \), then \( z = z' \). In other words, if \( \hat{F}(\tau, Z_t) \) is known for all \( \tau \geq 0 \), we can back out the value of \( Z_t \).
where we denote $\tilde{M}_t = (\tilde{M}_{Zt}, \tilde{M}_{Ut})$. Moreover, the state price density can be written

$$\zeta_t = e^{-\alpha t}(\hat{\phi} + \hat{\psi}^T Z_t).$$

(13)

Now, since the process $Z_t$ has a linear drift that depends only on $Z_t$ itself, and since the state price density also depends only on $Z_t$, we can view $Z_t$ as the factor process of an $m$-dimensional linear-rational term structure model (12)–(13), which is equivalent to (2)–(3). In view of Theorem 2.2, this leads to an interpretation of Theorem 2.4(iii): the model (12)–(13) is minimal in the sense that its own term structure kernel is trivial.

Carrying this observation further, we see that if the unspanned factors $U_t$ do not enter into the dynamics of $\tilde{M}_{Zt}$, then $Z_t$ is a fully autonomous Markov process, assuming that $X_t$ is Markovian. In this case $U_t$ would be redundant and play no role in the model. However, if $U_t$ does enter into the dynamics of $\tilde{M}_{Zt}$, then the unspanned factors would not be redundant. This situation is what gives rise to USV, and is discussed in Section 3.2.

The $(Z_t, U_t)$-coordinates fully reveal any unspanning property of the linear-rational term structure model. If our interest is in unspanned factors, why did we not specify the model in these coordinates in the first place? The reason is that we have to control the interplay of the factor dynamics with the state space. Note that $E$ has to lie in the half-space $\{\phi + \psi^T x > 0\}$, and thus always has boundaries. The invariance of $E$ with respect to the factor dynamics $X_t$ is a non-trivial property, which is much simpler to control if $E$ has some regular shape (below we will consider $E = \mathbb{R}^d_+$).

The shape of the transformed state space $\hat{E} = S(E)$ may be deformed, and it is essentially impossible to specify a priori conditions on the $(Z_t, U_t)$-dynamics (12) that would assert the invariance of $\hat{E}$.

Finally, note that even if the term structure kernel is trivial, $U = \{0\}$, the short end of the term structure may nonetheless be insensitive to movements of the state along certain directions. In view of Theorem 2.2, for $d \geq 3$ we can have $U = \{0\}$ while still there exists a non-zero vector $\xi$ such that $\psi^T \xi = \psi^T \kappa \xi = 0$. This implies that the short rate function is constant along $\xi$, see (6). On the other hand, we can still, in the generic case, recover $X_t$ from a snapshot of the term structure, see Theorem 2.5.

### 2.3 Swaps and Swaptions

The linear-rational term structure models have the important advantage of allowing for tractable swaption pricing.
A fixed versus floating interest rate swap is specified by a tenor structure of reset and payment dates $T_0 < T_1 < \cdots < T_n$, where we take $\Delta = T_i - T_{i-1}$ to be constant for simplicity, and a pre-determined annualized rate $K$. At each date $T_i$, $i = 1, \ldots, n$, the fixed leg pays $\Delta K$ and the floating leg pays LIBOR accrued over the preceding time period. From the perspective of the fixed-rate payer, the value of the swap at time $t \leq T_0$ is given by

$$
\Pi_t^{\text{swap}} = P(t, T_0) - P(t, T_n) - \Delta K \sum_{i=1}^{n} P(t, T_i).
$$

The time-$t$ forward swap rate, $S_{t}^{T_0, T_n}$, is the strike rate $K$ that makes the value of the swap equal to zero. It is given by

$$
S_{t}^{T_0, T_n} = \frac{P(t, T_0) - P(t, T_n)}{\sum_{i=1}^{n} \Delta P(t, T_i)}.
$$

The forward swap rate becomes the spot swap rate at time $T_0$.

A payer swaption is an option to enter into an interest rate swap, paying the fixed leg at a pre-determined rate and receiving the floating leg. A European payer swaption expiring at $T_0$ on a swap with the characteristics described above has a value at expiration of

$$
C_{T_0} = (\Pi_{T_0}^{\text{swap}})^+ = \left( \sum_{i=0}^{n} c_i P(T_0, T_i) \right)^+ = \frac{1}{\zeta_{T_0}} \left( \sum_{i=0}^{n} c_i \mathbb{E} [\zeta_{T_i} | \mathcal{F}_{T_0}] \right)^+,
$$

for coefficients $c_i$ that can easily be read off the expression (14).

In a linear-rational term structure model, the conditional expectations $\mathbb{E} [\zeta_{T_i} | \mathcal{F}_{T_0}]$ are linear functions of $X_{T_0}$, with coefficients that are explicitly given in terms of the model parameters, see Lemma A.1. Specifically, we have

$$
C_{T_0} = \frac{1}{\zeta_{T_0}} p_{\text{swap}}(X_{T_0})^+.
$$

---

10For expositional ease, we assume that the payments on the fixed and floating legs occur at the same frequency. In reality, in the USD market fixed-leg payments occur at a semi-annual frequency, while floating-leg payments occur at a quarterly frequency. However, only the frequency of the fixed-leg payments matter for the valuation of the swap.

11This valuation equation, which was the market standard until a few years ago, implicitly assumes that payments are discounted with a rate that incorporates the same credit and liquidity risk as LIBOR. In reality, swap contracts are virtually always collateralized, which makes swap (and swaption) valuation significantly more involved; see, e.g., Johannes and Sundaresan (2007) and Filipović and Trolle (2013). In the present paper we simplify matters by adhering to the formula (14).
where $p_{\text{swap}}$ is the explicit linear function

$$
p_{\text{swap}}(x) = \sum_{i=0}^{n} c_i e^{-\alpha T_i} \left( \phi + \psi^\top \theta + \psi^\top e^{-\kappa (T_i - T_0)} (x - \theta) \right).
$$

The swaption price at time $t \leq T_0$ is then obtained by an application of the fundamental pricing formula (1), which yields

$$
\Pi^{\text{swpt}}_t = \frac{1}{\zeta_t} \mathbb{E}[\zeta_{T_0} \mathbb{E}[C_{T_0} \mid \mathcal{F}_t]] = \frac{1}{\zeta_t} \mathbb{E} [p_{\text{swap}}(X_{T_0})^+ \mid \mathcal{F}_t].
$$

(16)

To compute the price one has to evaluate the conditional expectation on the right side of (16). If the conditional distribution of $X_{T_0}$ given $\mathcal{F}_t$ is known, this can be done via direct numerical integration over $\mathbb{R}^d$. This appears to be a challenging problem in general; fortunately there is an alternative approach based on Fourier transform methods that tends to perform better in practice.

**Theorem 2.6.** Define $\hat{q}(z) = \mathbb{E}[\exp (z p_{\text{swap}}(X_{T_0})) \mid \mathcal{F}_t]$ for every $z \in \mathbb{C}$ such that the conditional expectation is well-defined. Pick any $\mu > 0$ such that $\hat{q}(\mu) < \infty$.

Then the swaption price is given by

$$
\Pi^{\text{swpt}}_t = \frac{1}{\zeta_t} \frac{1}{\pi} \int_0^{\infty} \Re \left[ \frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda.
$$

Theorem 2.6 reduces the problem of computing an integral over $\mathbb{R}^d$ to that of computing a simple line integral. Of course, there is a price to pay: we now have to evaluate $\hat{q}(\mu + i\lambda)$ efficiently as $\lambda$ varies through $\mathbb{R}_+$. This problem can be approached in various ways depending on the specific class of factor processes under consideration. In our empirical evaluation we focus on square-root factor processes, for which computing $\hat{q}(z)$ amounts to solving a system of ordinary differential equations, see Section 4.2.

It is often more convenient to represent swaption prices in terms of implied volatilities. In the USD market, the market standard is the “normal” (or “absolute” or “basis point”) implied volatility, which is the volatility parameter that matches a given price when plugged into the pricing formula that assumes a normal distribution for the underlying forward swap rate.\footnote{Alternatively, a price may be represented in terms of “log-normal” (or “percentage”) implied volatility, which assumes a log-normal distribution for the underlying forward swap rate.}
the forward swap rate \( K = S_t^{T_b,T_n}, \) see (15)), there is a particularly simple relation between the swaption price and the normal implied volatility, \( \sigma_{N,t}, \) given by

\[
\Pi_t^{\text{swpt}} = \sqrt{T_0 - t - 1} \frac{1}{\sqrt{2\pi}} \left( \sum_{i=1}^{n} \Delta P(t, T_i) \right) \sigma_{N,t};
\]

(17)

see, e.g., Corp (2012).

2.4 Comparison with Other Models

When the factor process \( X_t \) is Markovian, the linear-rational framework falls in the broad class of models contained under the potential approach laid out in Rogers (1997). There the state price density is modeled by the expression

\[
\zeta_t = e^{-at} R_{\alpha} g(X_t),
\]

where \( R_{\alpha} \) is the resolvent operator corresponding to the Markov process \( X_t, \) and \( g \) is a suitable function. In our setting we would have \( R_{\alpha} g(x) = \phi + \psi^\top x, \) and thus \( g(x) = (\alpha - \mathcal{G}) R_{\alpha} g(x) = \alpha \phi - \psi^\top \kappa \theta + \psi^\top (\alpha + \kappa) x, \) where \( \mathcal{G} \) is the generator of \( X_t. \)

Another related setup which slightly pre-dates the potential approach is the framework of Flesaker and Hughston (1996). The state price density now takes the form

\[
\zeta_t = \int_{t}^{\infty} M_{tu} \mu(u) du,
\]

where for each \( u, (M_{tu})_{0 \leq t \leq u} \) is a martingale. The Flesaker-Hughston framework is related to the potential approach (and thus to the linear-rational framework) via the representation

\[
e^{-at} R_{\alpha} g(X_t) = \int_{t}^{\infty} \mathbb{E} \left[ e^{-au} g(X_u) \mid \mathcal{F}_t \right] du,
\]

which implies \( M_{tu} \mu(u) = \mathbb{E} [e^{-au} g(X_u) \mid \mathcal{F}_t]. \) The linear-rational framework fits into this template by taking \( \mu(u) = e^{-au} \) and \( M_{tu} = \mathbb{E} [g(X_u) \mid \mathcal{F}_t] = \alpha \phi + \alpha \psi^\top \theta + \psi^\top (\alpha + \kappa) e^{-\kappa (u-t)} (X_t - \theta), \) where \( g(x) = \alpha \phi - \psi^\top \kappa \theta + \psi^\top (\alpha + \kappa) x \) was chosen as above. One member of this class, introduced in Flesaker and Hughston (1996), is the one-factor rational log-normal model. The simplest time-homogeneous version of this model is, in the notation of (2)–(3), obtained by taking \( \phi \) and \( \psi \) positive, \( \kappa = \theta = 0, \) and letting the martingale part \( M_t \) of the factor process \( X_t \) be geometric Brownian motion.

Finally, a more recently introduced set of models that are closely related to those mentioned above is the linearity-generating family studied in Gabaix (2009) and
Carr, Gabaix, and Wu (2009). The model considered by Carr, Gabaix, and Wu (2009) falls within the linear-rational class: One sets $\phi = 0$, $\alpha = 0$, $\theta = 0$, and lets the martingale part $M_t$ of the factor process $X_t$ be given by

$$dM_t = e^{-\kappa t}\beta dN_t,$$

where $\beta$ is a vector in $\mathbb{R}^d$ and $N_t$ is an exponential martingale of the form

$$\frac{dN_t}{N_t} = \sum_{i=1}^{m} \sqrt{v_{it}}dB_{it},$$

for independent Brownian motions $B_{it}$ and processes $v_{it}$ following square-root dynamics. The factor process in this model is non-stationary due the time-inhomogeneous volatility specification. In fact, assuming the eigenvalues of $\kappa$ have positive real part (which is the case in Carr, Gabaix, and Wu (2009)), the volatility of $X_t$ tends to zero as time goes to infinity, and the state itself converges to zero almost surely. The models we consider in our empirical analysis are time-homogeneous and stationary. They also have a volatility structure that is very different from the specification in Carr, Gabaix, and Wu (2009).

A common feature of all the above models is that bond prices are given as a ratio of two functions of the state. This is of course an artifact of the form of the pricing equation (1), and the fact that the state price density is the primitive object that is being modeled.

### 3 Linear-Rational Diffusion Models

We now specialize the linear-rational framework (2)–(3) to the case where the factor process has diffusive dynamics of the form

$$dX_t = \kappa(\theta - X_t)dt + \sigma(X_t)dB_t.$$  \hspace{1cm} (18)

Here $\sigma : E \to \mathbb{R}^{d\times d}$ is measurable, and $B_t$ is $d$-dimensional Brownian motion. We denote the diffusion matrix by $a(x) = \sigma(x)\sigma(x)^\top$, and assume that it is differentiable. The goal of this section is two-fold: we first discuss how the notion of spanning in Section 2.2 relates to bond market completeness. Then, in the case where unspanned factors are present, we answer the question of when these unspanned factors give rise to USV.
3.1 Bond Markets

Bond price volatilities will be central to the discussion, so we begin by considering the dynamics of bond prices. To this end, first observe (via a short calculation using Itô’s formula) that the dynamics of the state price density can be written

$$\frac{d\zeta_t}{\zeta_t} = -r_t dt - \lambda_t^\top dB_t,$$

where the short rate \( r_t \) is given by (6), and \( \lambda_t = -\sigma(X_t)^\top \psi/(\phi + \psi^\top X_t) \) is the market price of risk. It then follows that the dynamics of \( P(t,T) \) is

$$\frac{dP(t,T)}{P(t,T)} = (r_t + \nu(t,T)^\top \lambda_t) dt + \nu(t,T)^\top dB_t,$$

where the volatility vector is given by

$$\nu(t,T) = \frac{\sigma(X_t)^\top \nabla F(T-t,X_t)}{F(T-t,X_t)}.$$

It is intuitively clear that a non-trivial term structure kernel gives rise to bond market incompleteness in the sense that not every contingent claim can be hedged using bonds. Conversely, one would expect that whenever the term structure kernel is trivial, bond markets are complete. In this section we confirm this intuition. The following definition of completeness is standard.

**Definition 3.1.** We say that bond markets are complete if for any \( T \geq 0 \) and any bounded \( \mathcal{F}_T \)-measurable random variable \( C_T \), there is a set of maturities \( T_1, \ldots, T_m \) and a self-financing trading strategy in the bonds \( P(t,T_1), \ldots, P(t,T_m) \) and the money market account, whose value at time \( T \) is equal to \( C_T \).

Our next result clarifies the connection between bond market completeness and the existence of unspanned factors. We assume that the filtration is generated by the Brownian motion, and that the volatility matrix of the factor process itself is almost surely invertible. Otherwise there would be measurable events which cannot be generated by the factor process, and bond market completeness would fail.

**Theorem 3.2.** Assume that the filtration \( \mathcal{F}_t \) is generated by the Brownian motion \( B_t \), that \( \sigma(X_t) \) is invertible \( dt \otimes d\mathbb{P} \)-almost surely, and that \( \phi + \psi^\top \theta \neq 0 \). Then the following conditions are equivalent:

(i) bond markets are complete;
(ii) $\text{span}\{\nabla F(\tau, X_t) : \tau \geq 0\} = \mathbb{R}^d$, $dt \otimes d\mathbb{P}$-almost surely;

(iii) $U = \{0\}$ and $\kappa$ is invertible;

(iv) The term structure $F(\tau, x)$ is injective.

\subsection*{3.2 Unspanned Stochastic Volatility Factors}

We now refine the discussion in Section 2.2 by singling out those unspanned factors that give rise to USV. To this end we describe directions $\xi \in \mathbb{R}^d$ with the property that movements of the state vector along $\xi$ influence neither the bond return volatilities, nor the covariations between returns on bonds with different maturities.

According to (19), the covariation at time $t$ between the returns on two bonds with maturities $T_1$ and $T_2$ is given by

$$\nu(t, T_1) \top \nu(t, T_2) = G(T_1 - t, T_2 - t, X_t),$$

where we define

$$G(\tau_1, \tau_2, x) = \frac{\nabla F(\tau_1, x) \top a(x) \nabla F(\tau_2, x)}{F(\tau_1, x) F(\tau_2, x)}. \quad (20)$$

In analogy with Definition 2.1 we introduce the following notion:

**Definition 3.3.** The variance-covariance kernel, denoted by $\mathcal{W}$, is given by

$$\mathcal{W} = \bigcap_{\tau_1, \tau_2 \geq 0} \ker G(\tau_1, \tau_2, \cdot).$$

That is, $\mathcal{W}$ consists of all $\xi \in \mathbb{R}^d$ such that $\nabla G(\tau_1, \tau_2, x) \top \xi = 0$ for all $\tau_1, \tau_2 \geq 0$ and all $x \in E$. We say that the model exhibits USV if there are elements of the term structure kernel that do not lie in the variance-covariance kernel—i.e., if $U \backslash \mathcal{W} \neq \emptyset$.

Analogously to Section 2.2 we may now transform the state space so that the intersection $U \cap \mathcal{W}$ of the term structure kernel and variance-covariance kernel corresponds to the last components of the state vector. To this end, let $S$ be an invertible linear transformation satisfying (11), with the additional property that

$$S(U \cap \mathcal{W}) = \{0\} \times \{0\} \times \mathbb{R}^q,$$

where $q = \dim U \cap \mathcal{W}$, and $p + q = n = \dim U$. The unspanned factors then decompose accordingly into $U_t = (V_t, W_t)$. Movements of $W_t$ affect neither the term structure, nor bond return volatilities or covariations. In contrast, movements of $V_t$, while having no effect on the term structure, do impact bond return volatilities or covariations. For this reason we refer to $V_t$ as USV factors, whereas $W_t$ is referred to as residual factors. Note that the residual factors $W_t$ may still have an indirect impact.
on the distribution of future bond prices. An example in Appendix B illustrates this fact.

Whether a given linear-rational term structure model exhibits USV depends on how $\sigma$ interacts with the other parameters of the model. Theorem A.7 in the appendix gives a description of the variance-covariance kernel $W$, which facilitates checking the presence of USV. As a corollary we obtain the following useful sufficient condition for USV. This condition is, for example, satisfied for the square root model discussed in Section 4. It is stated in terms of the diffusion matrix $\hat{a}(z,u)$ of the transformed factor process $\hat{X}_t = (Z_t, U_t)$, given by $\hat{a}(z,u) = S a(S^{-1}(z,u)) S^T$.

**Corollary 3.4.** Assume for every $j \in \{1, \ldots, n\}$, there exists $i \in \{1, \ldots, m\}$ such that $\hat{a}_{ii}(z,u)$ is not constant in $u_j$. Then $U \cap W = \{0\}$, and therefore every unspanned factor is in fact a USV factor.

## 4 The Linear-Rational Square-Root Model

The primary example of a linear-rational diffusion model (18) with state space $E = \mathbb{R}^d_+$ is the linear-rational square-root (LRSQ) model. It is based on a square-root factor process of the form

$$dX_t = \kappa(\theta - X_t)dt + \text{Diag}(\sigma_1 \sqrt{X_{1t}}, \ldots, \sigma_d \sqrt{X_{dt}}) \, dB_t,$$

(21)

with parameters $\sigma_i$. In this section we consider this model, focusing on how unspanned stochastic volatility can be incorporated, and how swaption pricing can be done efficiently. This lays the groundwork for our empirical analysis.

### 4.1 Unspanned Stochastic Volatility

The aim is now to construct a large class of LRSQ specifications with $m$ term structure factors and $n$ USV factors. Other constructions are possible, but the one given here is more than sufficient for the applications we are interested in.

As a first step we show that the LRSQ model admits a canonical representation.

---

13 The “unspanned components of the macro variables $M_t$” in Joslin, Priebsch, and Singleton (2010) and the “hidden factors $h_t$” in Duffee (2011), both of which are Gaussian exponential-affine models, are residual factors. Indeed, they neither show up in the bond yields nor in the bond volatilities. They enter through the equivalent change of measure from the risk-neutral $Q$ to the historical measure $P$, and thus affect the distribution of future bond prices.
Theorem 4.1. The short rate (6) is bounded from below if and only if, after a coordinatewise scaling of the factor process (21), we have $\zeta_t = e^{-\alpha t} (1 + \mathbf{1}^\top X_t)$. In this case, the extremal values in (7) are given by $\alpha^* = \max \mathcal{S}$ and $\alpha_* = \min \mathcal{S}$ where

$$
\mathcal{S} = \{ \mathbf{1}^\top \kappa \theta, -\mathbf{1}^\top \kappa_1, \ldots, -\mathbf{1}^\top \kappa_d \}.
$$

In accordance with this result, we always let the state price density be given by $\zeta_t = e^{-\alpha t} (1 + \mathbf{1}^\top X_t)$ when considering the LRSQ model.

Now fix nonnegative integers $m \geq n$ with $m + n = d$, representing the desired number of term structure and USV factors, respectively. We start from the observation that (11) holds if and only if the last $n$ column vectors of $S^{-1}$ form a basis of the term structure kernel $\mathcal{U}$. The first $m$ columns of $S^{-1}$ can be freely chosen, as long as all column vectors stay linearly independent. The next observation is that $\ker \mathbf{1}^\top$ is spanned by vectors of the form $-e_i + e_j$, for $i < j$, where $e_i$ denotes the $i$th standard basis vector in $\mathbb{R}^d$. We now deliberately choose $-e_i + e_{m+i}$, $i = 1, \ldots, n$, as basis for $\mathcal{U}$, which thus lies in $\ker \mathbf{1}^\top$ as required by Theorem 2.2. This amounts to specifying the invertible linear transformation $S$ on $\mathbb{R}^d$ by

$$
S = \begin{pmatrix} \text{Id}_m & A \\ 0 & \text{Id}_n \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} \text{Id}_m & -A \\ 0 & \text{Id}_n \end{pmatrix},
$$

where $A \in \mathbb{R}^{m \times n}$ is given by

$$
A = \begin{pmatrix} \text{Id}_n \\ 0 \end{pmatrix}.
$$

The parameters appearing in the description (21) of the factor process $X_t$ can now be specified with the aid of Theorem 2.4, and for this it is convenient to introduce the index sets $I = \{1, \ldots, m\}$ and $J = \{m+1, \ldots, d\}$. We write the mean reversion matrix $\kappa$ in block form as

$$
\kappa = \begin{pmatrix} \kappa_{II} & \kappa_{IJ} \\ \kappa_{JI} & \kappa_{JJ} \end{pmatrix},
$$

where $\kappa_{IJ}$ denotes the submatrix whose rows are indexed by $I$ and columns by $J$, and similarly for $\kappa_{II}$, $\kappa_{JI}$, $\kappa_{JJ}$. We require that $\kappa_{IJ}$ satisfy the restriction

$$
\kappa_{IJ} = \kappa_{II} A - A \kappa_{JJ} + A \kappa_{JI} A.
$$

(22)

The level of mean reversion is taken to be a vector $\theta = (\theta_I, \theta_J) \in \mathbb{R}^m \times \mathbb{R}^n$, and we fix some volatility parameters $\sigma_i > 0$, $i = 1, \ldots, d$. To guarantee that a solution to (21) exists, we impose the standard admissibility conditions that $\kappa \theta \in \mathbb{R}^d_+$ and the off-diagonal elements of $\kappa$ be nonpositive, see e.g. Filipović (2009, Theorem 10.2).
To confirm that this model indeed exhibits USV, we consider the dynamics of the transformed state vector $\hat{X}_t = SX_t = (Z_t, U_t)$. The transformed parameters are obtained from (9). Due to (22) and the form of $S$ we get

$$\hat{\kappa} = S\kappa S^{-1} = \begin{pmatrix} \kappa_{II} + A\kappa_{JI} & 0 \\ \kappa_{JI} & \kappa_{JJ} - \kappa_{JI}A \end{pmatrix}, \quad \hat{\theta} = \begin{pmatrix} \theta_I + A\theta_J \\ \theta_J \end{pmatrix}.$$  

The following result shows that this specification gives rise to at least $n$ USV factors.

**Corollary 4.2.** The dimension of the term structure kernel is at least $n$, $\dim U \geq n$, with equality if $\hat{\kappa}_{ZZ} = \kappa_{II} + A\kappa_{JI}$ satisfies Theorem 2.4(iii). In this case, if $\sigma_i \neq \sigma_{m+i}$ for $i = 1, \ldots, n$, then all the unspanned factors are in fact USV factors.

For our empirical analysis, we employ the following parsimonious specification that falls within the class described above.

**Definition 4.3.** The LRSQ($m, n$) specification is obtained by letting in the above construction $\kappa_{JI} = 0$ and $\kappa_{JJ} = A^\top \kappa_{II} A$ (this is the upper left $n \times n$ block of $\kappa_{II}$).

As an illustration, consider the LRSQ(1, 1) specification, where we have one term structure factor and one unspanned factor. It shows in particular that a linear-rational term structure model may exhibit USV even in the two-factor case.\(^{14}\)

**Example 4.4.** Under the LRSQ(1, 1) specification the mean reversion matrix is given by

$$\kappa = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{11} \end{pmatrix}.$$  

The term structure factor and unspanned factor thus become $Z_t = X_{1t} + X_{2t}$ and $U_t = X_{2t}$, respectively. The transformed mean reversion matrix $\hat{\kappa}$ coincides with $\kappa$,

$$\hat{\kappa} = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{11} \end{pmatrix},$$  

and the corresponding volatility matrix is

$$\hat{\sigma}(z, u) = \begin{pmatrix} \sigma_1 \sqrt{z_1 - u_1} & \sigma_2 \sqrt{u_1} \\ \sigma_2 \sqrt{u_1} & \sigma_2 \sqrt{u_1} \end{pmatrix}.$$  

Thus the transformed diffusion matrix $\hat{a}(z, u)$ satisfies $\hat{a}_{11}(z, u) = \sigma_1^2 z + (\sigma_2^2 - \sigma_1^2)u$. This is non-constant in $u$ as long as $\sigma_1 \neq \sigma_2$, as it should in view of Corollary 4.2. In particular, $U_t$ is a USV factor.

\(^{14}\)This contradicts the statement of Collin-Dufresne and Goldstein (2002, Proposition 3), which thus is incorrect.
Example 4.5. Consider now the LRSQ$(3,1)$ specification. In this case we have

$$\kappa = \begin{pmatrix}
\kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\
\kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{21} \\
\kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{31} \\
0 & 0 & 0 & \kappa_{11}
\end{pmatrix},$$

where it is straightforward to impose admissibility conditions on $\kappa$. The term structure factors and unspanned factors become

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \\ U_{1t} \end{pmatrix} = SX_t = \begin{pmatrix} X_{1t} + X_{4t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{pmatrix},$$

and the transformed mean reversion matrix is given by

$$\hat{\kappa} = \begin{pmatrix}
\kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\
\kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\
\kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\
0 & 0 & 0 & \kappa_{11}
\end{pmatrix}.$$

The corresponding volatility matrix is

$$\hat{\sigma}(z,u) = \begin{pmatrix}
\sigma_1 \sqrt{z_1 - u_1} & 0 & 0 & \sigma_4 \sqrt{u_1} \\
0 & \sigma_2 \sqrt{z_2} & 0 & 0 \\
0 & 0 & \sigma_3 \sqrt{z_3} & 0 \\
0 & 0 & 0 & \sigma_4 \sqrt{u_1}
\end{pmatrix}.$$

We now have $\hat{a}_{11}(z,u) = \sigma_1^2 z_1 + (\sigma_4^2 - \sigma_1^2) u_1$, which again demonstrates the presence of USV, provided $\sigma_1 \neq \sigma_4$.

4.2 Swaption Pricing

Swaption pricing becomes particularly tractable in the LRSQ model. Since $X_t$ is a square-root process, the function $\hat{q}(z)$ in Theorem 2.6 can be expressed using the exponential-affine transform formula that is available for such processes. Computing $\hat{q}(z)$ then amounts to solving a system of ordinary differential equations, which takes the following well-known form; see, e.g., Duffie, Pan, and Singleton (2000) and Filipović (2009, Theorem 10.3).
Lemma 4.6 (Exponential-Affine Transform Formula). For any $x \in \mathbb{R}^d_+, t \geq 0, u \in \mathbb{C}, v \in \mathbb{C}^d$ such that $E_x[|\exp(v^\top X_t)|] < \infty$ we have
\[ E_x[e^{u + v^\top X_t}] = e^{\Phi(t) + x^\top \Psi(t)}, \]
where $\Phi : \mathbb{R}_+ \to \mathbb{C}$, $\Psi : \mathbb{R}_+ \to \mathbb{C}^d$ solve the system
\[
\Phi'(\tau) = b^\top \Psi(\tau)
\Psi'_i(\tau) = \beta_i^\top \Psi(\tau) + \frac{1}{2}\sigma_i^2\Psi_i(\tau)^2, \quad i = 1, \ldots, d,
\]
with initial condition $\Phi(0) = u$, $\Psi(0) = v$. The solution to this system is unique.

In order for Theorem 2.6 to be applicable, it is necessary that some exponential moments of $p_{\text{swap}}(X_t)$ be finite. We therefore remark that for $X_t$ of the form (21), and for any $v \in \mathbb{R}^d$, $x \in \mathbb{R}^d_+, t \geq 0$, there is always some $\mu > 0$ (depending on $v, x, t$) such that $E_x[\exp(\mu v^\top X_t)] < \infty$. While it may be difficult a priori to decide how small $\mu$ should be, the choice is easy in practice since numerical methods diverge if $\mu$ is too large, resulting in easily detectable outliers.

5 Flexible Market Price of Risk Specification

The market price of risk in a linear-rational diffusion model is endogenous, and can be too restrictive to match certain empirical features of the data. In this section we describe a simple way of introducing flexibility in the market price of risk specification, which allows us to circumvent this issue.\textsuperscript{15}

The starting point is the observation that the interpretation of $\mathbb{P}$ as the historical probability measure is never used in the preceding sections. Indeed, the linear-rational framework can equally well be developed under some auxiliary probability measure $\mathbb{A}$ that is equivalent to $\mathbb{P}$. The state price density $\zeta_t \equiv \zeta_t^A = e^{-\alpha t}(\phi + \psi^\top X_t)$ and the martingale $M_t \equiv M_t^A$ are then understood with respect to $\mathbb{A}$. The factor process dynamics reads
\[ dX_t = \kappa(\theta - X_t)dt + dM_t^A, \]
and the basic pricing formula (1) becomes
\[ \Pi(t, T) = \frac{1}{\zeta_t^A} \mathbb{E}^A[\zeta_T^AC_T | \mathcal{F}_t]. \]
\textsuperscript{15}The same idea has been used in Cairns (2004, Chapter 8).
From this, using Bayes’ rule, we obtain the state price density with respect to $\mathbb{P}$, 

$$
\zeta_t^P = \zeta_t^A \frac{d\mathbb{A}}{d\mathbb{P}} \bigg|_{\mathcal{F}_t},
$$

so that

$$
\Pi(t, T) = \frac{1}{\zeta_t^P} \mathbb{E}_P \left[ \zeta_T^P C_T \mid \mathcal{F}_t \right].
$$

The bond prices $P(t, T) = F(T - t, X_t)$ are still given as functions of the factor process $X_t$, with the same $F(\tau, x)$ as in (5).

In the diffusion setup of Section 3, the martingale $M_t^A$ is now given by $dM_t^A = \sigma(X_t)dB_t^A$ for some $A$-Brownian motion $B_t^A$, and the market price of risk $\lambda_t \equiv \lambda_t^A = -\sigma(X_t)^\top \psi / (\phi + \psi^\top X_t)$ is understood with respect to $A$.

Having specified the model under the auxiliary measure $A$, we now have full freedom in choosing an equivalent change of measure from $A$ to the historical measure $\mathbb{P}$. Specifically, $\mathbb{P}$ can be defined using a density process of the form

$$
d\mathbb{P} \bigg|_{\mathcal{F}_t} = \exp \left( \int_0^t \delta_s^\top dB_s^A - \frac{1}{2} \int_0^t \|\delta_s\|^2 ds \right),
$$

for some appropriate Girsanov kernel $\delta_t$. The $\mathbb{P}$-dynamics of the factor process becomes

$$
dX_t = [\kappa(\theta - X_t) + \sigma(X_t)\delta_t]dt + \sigma(X_t)dB_t^P,
$$

for the $\mathbb{P}$-Brownian motion $dB_t^P = dB_t^A - \delta_t dt$. The state price density with respect to $\mathbb{P}$ follows the dynamics

$$
\frac{d\zeta_t^P}{\zeta_t^P} = -r_t dt - (\lambda_t^P)^\top dB_t^P,
$$

where the market price of risk $\lambda_t^P$ is now given by

$$
\lambda_t^P = \lambda_t^A + \delta_t = -\frac{\sigma(X_t)^\top \psi}{\phi + \psi^\top X_t} + \delta_t.
$$

Particular specifications of $\delta_t$ used in our empirical analysis are discussed in Section 6.

The volatility vector of the bond returns is invariant under an equivalent change of measure. Hence the covariation between the returns on two bonds with maturities $T_1$ and $T_2$, $\nu(t, T_1)^\top \nu(t, T_2) = G(T_1 - t, T_2 - t, X_t)$, is still given as function of the factor process, with the same $G(\tau_1, \tau_2, x)$ as in (20).
6 Data and Model Specifications

6.1 Swaps and Swaptions

The empirical analysis is based on a panel data set consisting of swaps and swaptions. At each observation date, we observe rates on spot-starting swap contracts with maturities of one, two, three, five, seven, and ten years, respectively. We also observe prices on swaptions with three-month option maturities, the same six swap maturities, and strikes equal to the forward swap rates. Such at-the-money-forward (ATMF) swaptions are the most liquid. We convert swaption prices into normal implied volatilities using (17) with zero-coupon bonds bootstrapped from the swap curve. The data is from Bloomberg and consists of composite quotes computed from quotes that Bloomberg collects from major banks and inter-dealer brokers. The sample period consists of 866 weekly observations from January 29, 1997 to August 28, 2013.

Table 1 shows summary statistics of swap rates (Panel A) and swaption IVs (Panel B). The term structure of swap rates is upward-sloping, on average, while the standard deviation of swap rates decreases with maturity. Time series of the 1-year, 5-year, and 10-year swap rate are displayed in Panel A1 of Figure 1. The 1-year swap rate fluctuates between a minimum of 0.32 percent (on October 17, 2012) and a maximum of 7.51 percent (on May 17, 2000), while the longer-term swap rates exhibit less variation. A principal component analysis (PCA) of weekly changes in swap rates shows that the first three factors explain 90, 7, and 2 percent, respectively, of the variation.

The term structure of swaption IVs is hump-shaped, on average, increasing from 81 bps at the 1-year swap maturity to 107 bps at the 7-year swap maturity. The standard deviation of swaption IVs is also a hump-shaped function of maturity. Time series of swaption IVs at the 1-year, 5-year, and 10-year swap maturities are displayed in Panel B1 of Figure 1. The swaption IV at the 1-year swap maturity fluctuates between a minimum of 17 bps (on October 10, 2012) and a maximum of 224 bps (on October 15, 2008), while swaption IVs at longer swap maturities fluctuate in a tighter range. Swaptions also display a high degree of commonality, with the first three factors from a PCA of weekly changes in swaption IVs explaining 87, 7, and 2 percent, respectively, of the variation.

6.2 Model Specifications

We restrict attention to the $LRSQ(m, n)$ specification, see Definition 4.3. We always set $m = 3$ (three term structure factors) and consider specifications with $n = 1$. 

(volatility of $Z_{1t}$ containing an unspanned component), $n = 2$ (volatility of $Z_{1t}$ and $Z_{2t}$ containing unspanned components), and $n = 3$ (volatility of all term structure factors containing unspanned components).

We price swaps and swaptions under $A$ and obtain the change of measure from $A$ to $P$ by specifying $\delta_t$ parsimoniously as

$$\delta_t = \left( \delta_1 \sqrt{X_{1t}}, \ldots, \delta_d \sqrt{X_{dt}} \right)^\top.$$

This choice is convenient as $X_t$ remains a square-root process under $P$, facilitating the use of standard estimation techniques from the vast body of literature on affine models. Specifically, we estimate by quasi-maximum likelihood in conjunction with Kalman filtering. Details are provided in Appendix C. Note that $X_t$ is not a square-root process under $Q$. This is in contrast to the exponential-affine framework.

In preliminary analyses, we found that the upper-triangular elements of $\kappa_{II}$ were always very close to zero. The same was true of $\kappa_{31}$. To obtain more parsimonious model specifications, we reestimate after setting to zero these elements of the mean-reversion matrix. The likelihood functions were virtually unaffected by this, so we henceforth study these constrained model specifications.

### 7 Results

#### 7.1 Maximum Likelihood Estimates

Table 2 displays parameter estimates and their asymptotic standard errors. A robust feature across all specifications is that the drift parameters align such that to a close approximation

$$\alpha = 1^\top \hat{\kappa}_{ZZ} \hat{\theta}_Z = -1^\top \hat{\kappa}_{ZZ,1} = -1^\top \hat{\kappa}_{ZZ,2} > -1^\top \hat{\kappa}_{ZZ,3} = \alpha_*. $$

Since $\hat{\kappa}_{ZZ}$ is lower triangular, $1^\top \hat{\kappa}_{ZZ,3} = \kappa_{33}$ and the expression for the short rate reduces to

$$r_t = (\alpha + \kappa_{33}) \frac{Z_{3t}}{1 + 1^\top Z_t}.$$ 

It is immediately clear that the range of $r_t$ is given by $r_t \in [0, \alpha + \kappa_{33})$; the lower bound is attained when $Z_{3t} = 0$, while the upper bound is not attainable. The table reports the upper bound on the short rate, which lies in a range from 22.4% to 125.3% across specifications.\(^{16}\) However, simulations show that the likelihood of

---

\(^{16}\)Incidentally, 22.4% is the historical maximum for the effective federal funds rate reached on July 22, 1981, during the monetary policy experiment.
observing short rates above 20% is negligible in all specifications. In contrast, there is a significant likelihood of observing very low short rates; see below. Since all model specifications are stationary (all eigenvalues of $\hat{\kappa}_{ZZ}$ are positive), $\alpha$ equals the infinite-maturity forward rate. This rate lies in a range from 5.37% to 6.27% across specifications, which appears reasonable.

From a practical perspective, the simple structure for the short rate gives the model flexibility to capture significant variation in longer-term rates during the latter part of the sample period, when policy rates were effectively zero. The reason is that a low value of $Z_{3t}$ constrains the short rate to be close to zero, which allows $Z_{1t}$ and $Z_{2t}$ the freedom to affect longer-term rates without having much impact on the short rate.

For the diffusion term, recall that if the $i$'th term structure factor exhibits USV, its instantaneous volatility is given by $\sqrt{\sigma_i^2 Z_{it} + (\sigma_i^2 + 3 - \sigma_i^2)U_{it}}$. We always have $\sigma_{i+3} > \sigma_i$ implying that the volatilities of the term structure factors are increasing in the USV factors.

### 7.2 Factors

Figure 2 displays the estimated factors. The first, second, and third column shows the factors of the $LRSQ(3,1)$, $LRSQ(3,2)$, and $LRSQ(3,3)$ model specification, respectively. The first, second, and third row shows $(Z_{1t}, U_{1t})$, $(Z_{2t}, U_{2t})$, and $(Z_{3t}, U_{3t})$, respectively. The factors are highly correlated across specifications, which is indicative of a stable factor structure. The USV factors are occasionally large relative to the term structure factors which gives the first indication of the importance of allowing for USV.

To better understand the factor dynamics, Figure 3 plots the instantaneous volatility of each term structure factor against its level. Again, the first, second, and third column corresponds to the $LRSQ(3,1)$, $LRSQ(3,2)$, and $LRSQ(3,3)$ specification, respectively, while the first, second, and third row corresponds to $Z_{1t}$, $Z_{2t}$, and $Z_{3t}$, respectively. The grey areas mark the possible range of factor volatilities, which is given by $\sigma_i \sqrt{Z_{it}}$ to $\sigma_{i+3} \sqrt{Z_{it}}$ in case the $i$'th term structure factor exhibits USV. Whenever USV is allowed, there appears to be significant variation in factor volatilities that is unrelated to the factor levels.

### 7.3 Specification Analysis

For each of the model specifications, we compute the fitted swap rates and swap option IVs based on the filtered state variables. We then compute weekly root mean

---

24
squared pricing errors (RMSEs) separately for swap rates and swaption IVs, thereby constructing two time series of RMSEs. The first three rows in Table 3 report the sample means of the RMSE time series for the three specifications. To investigate the performance of the model when policy rates are close to the zero lower bound (ZLB), we also split the sample period into a ZLB period and a pre-ZLB period. The beginning of the ZLB period is taken to be December 16, 2008, when the Federal Reserve reduced the federal funds rate from one percent to a target range of 0 to 1/4 percent. The next two rows report the mean difference in RMSEs between model specifications along with the associated $t$-statistics corrected for heteroscedasticity and serial correlation in parenthesis.

Even the most parsimonious $LRSQ(3,1)$ specification has a reasonable fit to the data. For instance, for the full sample period, the mean RMSEs for swap rates and swaption IVs are 5.71 bps and 7.21 bps, respectively. Adding one more USV factor decreases the mean RMSEs by 2.50 bps and 1.07 bps, respectively, which is both economically important and statistically significant. Adding an additional USV factor has a negligible impact on the mean RMSE for swap rates, but decreases the mean RMSE for swaption IVs by a further 1.93 bps, which again is both economically important and statistically significant.

Comparing across sub-samples, for all the specifications the fit is better in the pre-ZLB period than in the ZLB period, particularly for swap rates. Nevertheless, even during the ZLB period, the model performs well. For instance, in case of the $LRSQ(3,3)$ specification, the mean RMSEs for swap rates and swaption IVs are 4.89 bps and 4.68 bps, respectively, during the ZLB period, compared with 2.65 bps and 4.01 bps, respectively, during the pre-ZLB period. The performance of the $LRSQ(3,3)$ specification over time is illustrated in Figure 1 which shows the fitted time series of selected swap rates (Panel A2) and swaption IVs (Panel B2) as well as time series of the RMSEs for swap rates (Panel A3) and swaption IVs (Panel B3). The online appendix contains additional information about the fit of the specifications.

### 7.4 Risk Premium Dynamics

Swap returns can be computed in several ways. Here, we work with excess returns on forward-starting swaps, which are most easily computed from the available swap data. In the online appendix, we show that very similar results are obtained with excess returns on spot-starting swaps as well as zero-coupon bonds bootstrapped from swap rates.

Consider a strategy of entering into the forward-starting swap contract described
in Section 2.3 at time \( t < T_0 \), paying a fixed rate equal to the forward swap rate \( S_{T_0}^{T_n} \) and receiving floating. Assume also that an amount of capital \( C_t \) (covering at least the required margins) is allocated to the trade and earns the risk-free rate. At time \( T_0 \), the value of the swap is

\[
\Pi_{T_0}^{\text{swap}} = - \left( S_{T_0}^{T_n} - S_t^{T_n} \right) \left( \sum_{i=1}^{n} \Delta P(T_0, T_i) \right).
\]

Since the swap has zero initial value, the return on the strategy from \( t \) to \( T_0 \) in excess of the risk-free rate is

\[
R_{t,T_0}^e = \frac{\Pi_{T_0}^{\text{swap}}}{C_t}.
\]

We consider nonoverlapping monthly excess returns computed using closing prices on the last business day of each month (we have verified that our results hold true regardless of the day of the month that the strategy is initiated). We also assume that \( C_t \) equals the notional of the swap (equal to one), which corresponds to a “fully collateralized” swap position.\(^{17}\)

Table 4 shows unconditional mean and volatility of excess returns in addition to the unconditional Sharpe ratio. All statistics are annualized. Both the mean and volatility increases with swap maturity, while the Sharpe ratio decreases with swap maturity. A similar pattern for Sharpe ratios has been observed by Duffee (2010) and Frazzini and Pedersen (2014) in the case of Treasury bonds. The mean excess returns and Sharpe ratios are high in our sample, but are inflated by the downward trend in interest rates over the sample period.

We also consider conditional expected excess returns on swaps. Table 5 reports results from regressing nonoverlapping monthly excess swap returns on previous month’s term structure slope and implied volatility (including a constant); i.e.,

\[
R_{t,T_0}^e = \beta_0 + \beta_{\text{Slp}} S_{\text{Slp}} t + \beta_{\text{Vol}} V_{\text{Vol}} t + \epsilon_{t,T_0},
\]

where \( S_{\text{Slp}} t \) is the difference between the \((T_n - T_0)\)-year swap rate and 3-month LIBOR and \( V_{\text{Vol}} t \) the normal implied volatility for a swaption with a 3-month option expiry and a \((T_n - T_0)\)-year swap maturity. Excess returns are in percent, and both \( S_{\text{Slp}} t \) and \( V_{\text{Vol}} t \) are standardized to facilitate comparison of regression coefficients. Many papers, typically using long samples of Treasury data, find that the slope of the term structure predicts excess bond returns with a positive sign; see, e.g., Campbell and Shiller (1991) and Dai and Singleton (2002). In our more recent sample of

\(^{17}\)Alternatively, one could choose \( C_t \) to achieve a certain target volatility for excess returns as in Duarte, Longstaff, and Yu (2007).
swap data, the predictive power of the slope of the term structure is much weaker. Indeed, the regression coefficient is never statistically significant and is even negative for short swap maturities. In contrast, the predictive power of implied volatility is much stronger. The regression coefficient is statistically significant for short swap maturities and is positive for all swap maturities indicating a positive risk-return tradeoff in the swap market. Economically, volatility is also a more important predictor of excess returns. Taking the 5-year swap maturity as an example, a one standard deviation increase in volatility (the term structure slope) increases the one-month expected excess returns by 18 bps (6 bps) which should be put in relation to the unconditional mean one-month excess return of 28 bps.

We investigate if our model is able to capture these patterns for risk premia. We focus on the population properties of the model as this is much more demanding than using the fitted data. To this end, we simulate 50,000 years of monthly data (600,000 observations) from each of the three model specifications. The results are displayed in the lower panels in Tables 4 and 5. Table 4 shows that all specifications capture the pattern that the mean and volatility of excess returns increase with swap maturity, while the Sharpe ratio decreases with swap maturity (there is a small hump in the Sharpe ratio term structure for the $LRSQ(3,3)$ specification). The mean excess returns and Sharpe ratios are only about half of those observed in the data, which is not surprising given that we are simulating stationary samples of swap rates, while the swap rates in the data exhibit a downward trend over the sample period.

Table 5 shows that the model qualitatively captures the predictive power of implied volatility for excess returns. The size of the regression coefficients as well as the $R^2$'s increase with the number of USV factors. For the $LRSQ(3,3)$ specification, the amount of predictability in the data is about 60% of that in the data, which seems respectable considering the parsimonious market price of risk specification that we employ.

---

18For the Treasury market, there is mixed evidence for a risk-return tradeoff. Engle, Lilien, and Robins (1987) find evidence of a positive risk-return tradeoff in an ARCH-in-mean framework, while Duffee (2002) runs regressions similar to Table 5 and finds that volatility only weakly predicts return. The discrepancies between the results of Duffee (2002) and our results may be due to some combination of differences in sample periods (his sample ends where our sample begins), his use of historical volatilities vs. our use of forward-looking implied volatilities, and structural differences between the Treasury and swap markets. Corroborating our results, Kim and Singleton (2012) find that volatility predicts excess bond returns in Japan during the Japanese ZLB period.

19The processes are simulated using the “full truncation” Euler discretization scheme of Lord, Koekkoek, and van Dijk (2010). To minimize the discretization bias, we use 10 steps per day.
7.5 Term Structure Dynamics at the Zero Lower Bound

A key characteristic of the recent history of U.S. interest rates is the extended period of near-zero policy rates. Japan has experienced near-zero policy rates since the early 2000s. A challenge for term structure models is the ability to generate such extended periods of low short rates. A related challenge, as emphasized by Kim and Singleton (2012), is that close to the ZLB, the distribution of future short rates becomes highly asymmetric with the most likely (modal) value being significantly lower than the mean value.

To investigate the performance of our model along this dimension, we simulate 50,000 years of weekly data (2,600,000 observations) from the three model specifications. Taking the $LRSQ(3,1)$ specification as an example and conditioning on the short rate being between 0 and 25 basis points, Panels A-C in Figure 4 show the frequency distribution of the future short rate at a 1-year, 2-year, and 5-year horizon, respectively. Clearly the model generates persistently low rates; conditional on the short rate being between 0 and 25 basis points at a given point in time, the likelihood of the short rate being in the same interval in two year’s time is more than 50%. Panel D shows the mean and modal paths of the short rate. Indeed, the model generates a significant spread between the mean and modal paths, with the mean value rising to 1.21% at a 5-year horizon but the most likely value remaining at 0.125% (the midpoint of the lowest rate interval).

Another consequence of extended periods of low short rates is a changing characteristic of the factor loadings, particularly the “level” factor. Table 6 shows the factor loadings of the first principal component of the term structure during the pre-ZLB period (Panel A) and during the ZLB period (Panel B). The factor loadings are constructed from the covariance matrix of weekly changes in swap rates. During the pre-ZLB period, the loadings are relatively flat (in fact decreasing slightly with maturity) giving rise to the notion of a “level” factor. However, during the ZLB period, the loadings increase strongly with maturity rising from 0.10 for the 1-year maturity to 0.50 for the 10-year maturity. As such, the first factor effectively becomes a “slope” factor.

The table also reports the factor loadings in the simulated weekly data, where the ZLB sample consists of those observations where 3-month LIBOR is less than 50 basis points. For all the specifications, the loadings are flat during the non-ZLB periods and strongly increasing during ZLB periods. In case of the $LRSQ(3,3)$ specification, the loadings during ZLB periods are very close to those observed in the data. As such, in contemporaneous work, Kim and Priebsch (2013) also note the changing characteristics of factor loadings during the ZLB period.

\footnote{In contemporaneous work, Kim and Priebsch (2013) also note the changing characteristics of factor loadings during the ZLB period.}
such, the model captures how the level factor morphs into a slope factor when short rates are near the ZLB.

7.6 Volatility Dynamics at the Zero Lower Bound

A large literature has investigated the dynamics of interest rate volatility. A particular focus has been on the extent to which variation in volatility is related to variation in the term structure, with most papers finding that a significant component of volatility is only weakly related to term structure movements.\footnote{See Collin-Dufresne and Goldstein (2002) and subsequent papers by Heidari and Wu (2003), Andersen and Benzoni (2010), Li and Zhao (2006), Li and Zhao (2009), Trolle and Schwartz (2009), and Collin-Dufresne, Goldstein, and Jones (2009), among others. The issue is not without controversy, however, with Fan, Gupta, and Ritchken (2003), Jacobs and Karoui (2009), and Bikbov and Chernov (2009) providing a sceptical appraisal of the evidence.} We first ask if our model can replicate the degree of USV observed in the data. Table 7 reports $R^2$s from regressing weekly changes in normal implied swaption volatilities on the first three principal components (PCs) of weekly changes swap rates. Squared PCs and a constant are also included in the regressions. This gives a rough idea about the fraction of volatility risk that can be hedged by trading in swaps. In the data, the $R^2$s range from 0.11 to 0.20. In the simulated weekly data, the $LRSQ(3,1)$ specification generates too low a degree of USV with $R^2$s in a range from 0.17 to 0.61. In contrast, both the $LRSQ(3,2)$ and $LRSQ(3,3)$ specifications generate $R^2$s that are close to those observed in the data.

Next, we look more closely at the relation between volatility and the level of rates. We focus on the volatility of the 1-year swap rate, since this is the rate that is nearest to the ZLB during the sample period. Figure 5 shows the swaption IV of the 1-year swap maturity (in basis points) plotted against the 1-year swap rate. It strongly indicates that volatility becomes more level-dependent as the underlying interest rate approaches the ZLB.

To investigate the issue more formally, we regress weekly changes in the swaption IV at the 1-year swap maturity on weekly changes in the 1-year swap rate (including a constant); i.e.,

$$\Delta \sigma_{N,t} = \beta_0 + \beta_1 \Delta S_t + \epsilon_t.$$  

Result are displayed in the upper part of Table 8, with Newey and West (1987) $t$-statistics using four lags in parentheses. The first column shows results using the entire sample period. $\beta_1$ is positive and statistically significant ($t$-statistic of 2.80); however, the $R^2$ is small at 0.05.\footnote{Trolle and Schwartz (2013) also document a positive level-dependence in swaption IVs.} That is, unconditionally, there is a relatively
small degree of level-dependence in volatility.

The second to sixth column shows results conditional on the 1-year swap rate being in the intervals 0-0.01, 0.01-0.02, 0.02-0.03, 0.03-0.04, and 0.04-0.08, respectively. A clear pattern emerges. At low interest rates, $\beta_1$ is positive and highly statistically significant (t-statistic of 7.52), and the $R^2$ is very high at 0.52. In other words, there is a strong and positive relation between volatility and rate changes, when rates are close to the ZLB.\textsuperscript{23} However, as interest rates increase, the relation between volatility and rate changes becomes progressively weaker. Both $\beta_1$ and $R^2$ decrease, and when the 1-year swap rate is above 0.03, the $R^2$ is essentially zero.

We then perform the same analysis on the simulated weekly data. In terms of $R^2$s, the $LRSQ(3,1)$ specification matches the pattern in the data quite closely with an $R^2$ the same as that in the data, when rates are close to zero, and a fast decay in the $R^2$ as rates increase. Both the $LRSQ(3,2)$ and $LRSQ(3,3)$ specifications generate too low an $R^2$ at low rates and too slow a decay as rates increase. In terms of the regression coefficient, the specifications do not quite match the degree of level-dependence, but do match the decay in the regression coefficient as rates increase.

Figure 6 is an illustration of the performance of the $LRSQ(3,2)$ specification in capturing the changing level dependence. It shows weekly changes in the normal implied volatility of the 3-month option on the 1-year swap rate (in basis points) against weekly changes in the 1-year swap rate, conditional on the 1-year swap rate being in the intervals 0-0.01 (Panel A), 0.01-0.02 (Panel B), 0.02-0.03 (Panel C), and 0.03-0.04 (Panel D). Black crosses denote the data and grey dots denote the simulated data. Clearly, the model goes a long way towards matching the data.

8 Conclusion

We introduce the class of linear-rational term structure models, where the state price density is modeled such that bond prices become linear-rational functions of the current state. This class is highly tractable with several distinct advantages: i) ensures nonnegative interest rates, ii) easily accommodates unspanned factors as earlier literature has estimated generalized diffusion models for the short-term interest rate; see, e.g., Chan, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996), Conley, Hansen, Luttmer, and Scheinkman (1997), and Stanton (1997). These papers generally find a relatively strong level-dependence in interest rate volatility. However, much of this level-dependence can be attributed to the 1979-1982 monetary policy experiment, which is not representative of the current monetary policy regime.

\textsuperscript{23}In Japanese data, Kim and Singleton (2012) also note the high degree of level-dependence in volatility, when rates are close to the ZLB.
fecting volatility and risk premia, and iii) has analytical solutions to swaptions. A parsimonious specification of the model with three term structure factors and at least two unspanned factors has a very good fit to both interest rate swaps and swaptions since 1997. In particular, the model captures well the dynamics of risk premia as well as the dynamics of the term structure and volatility during the recent period of near-zero interest rates.

A Proofs

This appendix provides all proofs and additional auxiliary results needed for these proofs.

Proof of Formula (4)
The following lemma directly implies formula (4). It is also used in the proof of Lemma A.9.

Lemma A.1. Assume that $X$ is of the form (2) with integrable starting point $X_0$. Then for any bounded stopping time $\rho$ and any deterministic $\tau \geq 0$, the random variable $X_{\rho+\tau}$ is integrable, and we have

$$E[X_{\rho+\tau} | \mathcal{F}_\rho] = \theta + e^{-\kappa \tau}(X_\rho - \theta).$$

Proof. We first prove the result for $\rho = 0$. An application of Itô’s formula shows that the process

$$Y_t = \theta + e^{-\kappa(T-t)}(X_t - \theta)$$

satisfies $dY_t = e^{-\kappa(T-t)}dM_t$, and hence is a local martingale. It is in fact a true martingale. Indeed, integration by parts yields

$$Y_t = Y_0 + e^{-\kappa(T-t)}M_t - \int_0^t M_s \kappa e^{-\kappa(T-s)}ds,$$

from which the integrability of $X_0$ and $L^1$-boundedness of the martingale $M$ imply that $Y$ is bounded in $L^1$. Fubini’s theorem then yields, for any $0 \leq t \leq u$,

$$E[Y_u | \mathcal{F}_t] = Y_t + M_t \left[ e^{-\kappa(T-u)} - e^{-\kappa(T-t)} - \int_t^u \kappa e^{-\kappa(T-s)}ds \right]$$

$$= Y_t, \quad 31$$
showing that $Y$ is a true martingale. Since $Y_\tau = X_\tau$, it follows that
\[
\mathbb{E}[X_\tau \mid F_0] = Y_0 = \theta + e^{-\kappa \tau}(X_0 - \theta),
\]
as claimed. If $\rho$ is a bounded stopping time, then the $L^1$-boundedness of $Y$, and hence of $X$, implies that $X_\rho$ is integrable. The result then follows by applying the $\rho = 0$ case to the process $(X_{\rho+s})_{s \geq 0}$ and filtration $(F_{\rho+s})_{s \geq 0}$.

\begin{proof}[Proof of Theorem 2.2]
Observe, by taking the orthogonal complement in (8), that we must prove $U^\perp = \text{span}\{(\kappa^\top)^p \psi : p = 0, \ldots, d - 1\}$. By the Cayley-Hamilton theorem (see Horn and Johnson (1990, Theorem 2.4.2)) we may equivalently let $p$ range over all nonnegative integers. In other words, we need to prove
\[
\text{span}\{\nabla F(\tau, x) : \tau \geq 0, x \in E\} = \text{span}\{(\kappa^\top)^p \psi : p \geq 0\}.
\]
(23)
Denote the left side by $S$. A direct computation shows that the gradient of $F$ is given by
\[
\nabla F(\tau, x) = \frac{e^{-\alpha \tau}}{\phi + \psi^\top x}\left[e^{-\kappa^\top \tau} \psi - e^{\alpha \tau} F(\tau, x) \psi\right],
\]
whence $S = \text{span}\{e^{-\kappa^\top \tau} \psi - e^{\alpha \tau} F(\tau, x) \psi : \tau \geq 0, x \in E\}$. By the non-triviality assumption there are $x, y \in E$ and $\tau \geq 0$ such that $F(\tau, x) \neq F(\tau, y)$. It follows that $e^{\alpha \tau}(F(\tau, x) - F(\tau, y)) \psi$, and hence $\psi$ itself, lies in $S$. We deduce that $S = \text{span}\{e^{-\kappa^\top \tau} \psi : \tau \geq 0\}$, which coincides with the right side of (23). This proves the formal expression (8).

It remains to show that $U$ given by (8) equals $U'$, defined as the largest subspace of $\ker \psi^\top$ that is invariant under $\kappa$. It follows from (8), for $p = 0$, that $U$ is a subspace of $\ker \psi^\top$. Moreover, $U$ is invariant under $\kappa$: let $\xi \in U$. Then $\kappa \xi \in \ker \psi^\top \kappa^{p-1}$ for all $p \geq 1$, and hence $\kappa \xi \in U$. Since $U'$ is the largest subspace of $\psi^\top$ with this property, we conclude that $U \subseteq U'$. Conversely, let $\xi \in U'$. By invariance we have $\kappa^p \xi \in U' \subseteq \ker \psi^\top$, and hence $\xi \in \ker \psi^\top \kappa^p$, for any $p \geq 0$. Hence $\xi \in U$, and we conclude that $U' \subseteq U$.

\begin{proof}[Proof of Corollary 2.3]
Write $\Lambda = \text{Diag}(\lambda_1, \ldots, \lambda_d)$ and consider the matrix
\[
A = [\psi \kappa^\top \psi \cdots (\kappa^\top)^{d-1} \psi].
\]
Writing $\hat{\psi} = S^{-T}\psi$, the determinant of $A$ is given by

$$
\det A = \det (S^T) \det \left( \begin{array}{cccc}
\hat{\psi} & \Lambda \hat{\psi} & \cdots & \Lambda^{d-1} \hat{\psi}
\end{array} \right)
$$

$$
= \det (S^T) \hat{\psi}_1 \cdots \hat{\psi}_d \det \left( \begin{array}{cccc}
1 & \lambda_1 & \cdots & \lambda_{d-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \lambda_d & \cdots & \lambda_{d-1}
\end{array} \right)
$$

$$
= \det (S^T) \hat{\psi}_1 \cdots \hat{\psi}_d \prod_{1 \leq i < j \leq d} (\lambda_j - \lambda_i),
$$

where the last equality uses the formula for the determinant of the Vandermonde matrix. Theorem 2.2 now shows that the term structure kernel is trivial precisely when all eigenvalues of $\kappa$ are distinct and all components of $\hat{\psi}$ are nonzero, as was to be shown.

**Proof of Theorem 2.4**

The proof of the sufficiency direction of Theorem 2.4 will build on the following lemma. It provides a useful condition that guarantees the existence of a minimal number of unspanned factors.

**Lemma A.2.** Let $m, n \geq 0$ be integers with $m + n = d$. If the transformed model parameters (9)-(10) satisfy (i)-(ii) in Theorem 2.4, we have $S(U) \supseteq \{0\} \times \mathbb{R}^n$. In this case, we have $\dim U \geq n$.

**Proof.** Let $\{e_1, \ldots, e_d\}$ denote the canonical basis of $\mathbb{R}^d$. We have for $p = 0, \ldots, d-1$,

$$
\psi^T \kappa^p S^{-1} e_i = \hat{\psi}^T \hat{\kappa}^p e_i.
$$

Note that $\hat{\kappa}^p$ has the same block triangular structure as $\hat{\kappa}$ for all $p \geq 1$. Thus for $i = m + 1, \ldots, d$, the right side above is zero for all $p \geq 0$, so $S^{-1} e_i \in U$. We deduce that $S(U) \supseteq \{0\} \times \mathbb{R}^n$ holds, as claimed. \qed

We can now proceed to the proof of Theorem 2.4. The case $n = 0$ is immediate, so we consider the case $n \geq 1$. We write $\hat{U} = S(U)$ and assume (11) holds. That is, we have

$$
\hat{U} = \{0\} \times \mathbb{R}^n \subseteq \mathbb{R}^m \times \mathbb{R}^n.
$$

On the other hand, Theorem 2.2 yields

$$
\hat{U} = \left\{ \hat{\xi} \in \mathbb{R}^d : \hat{\psi}^T \hat{\kappa}^p \hat{\xi} = 0 \text{ for } p = 0, 1, \ldots, d-1 \right\}.
$$
Partition \( \hat{\psi} \) and \( \hat{\kappa} \) to conform with the product structure of \( \mathbb{R}^m \times \mathbb{R}^n \):

\[
\hat{\psi} = \left( \begin{array}{c} \hat{\psi}_Z \\ \hat{\psi}_U \end{array} \right) \in \mathbb{R}^{m+n}, \quad \hat{\kappa} = \left( \begin{array}{c} \hat{\kappa}_{ZZ} \\ \hat{\kappa}_{UZ} \\ \hat{\kappa}_{UU} \end{array} \right) \in \mathbb{R}^{(m+n) \times (m+n)}.
\]

It then follows from (25) and (26) that we have \( \hat{\psi}^\top \hat{\kappa} \xi = 0 \) for any \( \hat{\xi}_U \in \mathbb{R}^n \). Hence \( \hat{\psi}_U = 0 \), which proves (i). As a consequence, we have

\[
\hat{\psi}^\top \hat{\kappa} = \left( \begin{array}{c} \hat{\psi}^\top \hat{\kappa}_{ZZ} \\ \hat{\psi}^\top \hat{\kappa}_{UZ} \\ \hat{\psi}^\top \hat{\kappa}_{UU} \end{array} \right)
\]

for \( p = 1 \). Suppose we know (27) holds for some \( p \geq 1 \). Then (25) and (26) imply, for any \( \hat{\xi}_U \in \mathbb{R}^n \),

\[
0 = \hat{\psi}^\top \hat{\kappa} \left( \begin{array}{c} 0 \\ \xi_U \end{array} \right) = \hat{\psi}^\top \hat{\kappa}_{ZZ} \hat{\kappa}_{UZ} \xi_U,
\]

and hence

\[
\hat{\psi}^\top \hat{\kappa}_{ZZ} \hat{\psi} = 0.
\]

Multiplying both sides of (27) by \( \hat{\kappa} \) from the right then shows that (27) holds also for \( p + 1 \). It follows by induction that (27) and (28) hold for all \( p \geq 1 \).

Now, pick any \( \hat{\xi}_Z \) such that \( \hat{\psi}^\top \hat{\kappa}_Z \hat{\xi}_Z = 0 \) for all \( p \geq 0 \). Then the vector \( \hat{\xi} = (\hat{\xi}_Z, 0) \in \mathbb{R}^m \times \mathbb{R}^n \) satisfies \( \hat{\psi}^\top \hat{\kappa}_p \hat{\xi} = 0 \) for all \( p \geq 0 \). Hence \( \hat{\xi} \in \hat{U} \) by (26), and then \( \hat{\xi}_Z = 0 \) by (25). This proves (iii). Finally, since (28) holds for all \( p \geq 1 \), the range of \( \hat{\kappa}_{ZU} \) lies in the kernel of \( \hat{\psi}^\top \hat{\kappa}_{ZZ} \) for all \( p \geq 1 \). But by (iii) this implies that the range of \( \hat{\kappa}_{ZU} \) consists of the zero vector, which is to say that \( \hat{\kappa}_{ZU} = 0 \). This proves (ii).

We now prove the converse part of the theorem, and assume that \( \hat{\kappa} \) and \( \hat{\psi} \) given in (9) and (10) satisfy (i)–(iii). Lemma A.2 implies that \( S(\hat{U}) \supseteq \{0\} \times \mathbb{R}^n \). Consider now a vector \( \xi \) in the span of \( S^{-1} e_1, \ldots, S^{-1} e_m \), where \( \{e_1, \ldots, e_d\} \) denotes the canonical basis of \( \mathbb{R}^d \). Then \( S \xi = (\hat{\xi}_Z, 0) \in \mathbb{R}^m \times \mathbb{R}^n \). By (i) and (ii) we have

\[
\hat{\psi}^\top \hat{\kappa}_p \xi = \hat{\psi}^\top \hat{\kappa}_p S \xi = \hat{\psi}^\top \hat{\kappa}_{ZZ} \hat{\xi}_Z.
\]

If \( \xi \in \hat{U} \), the left side is zero for all \( p = 0, \ldots, d - 1 \). Hence so is the right side, which by (iii) implies \( \xi = 0 \). We deduce that \( \hat{U} = \text{span}\{S^{-1} e_{m+1}, \ldots, S^{-1} e_d\} \) and hence \( S(\hat{U}) = \{0\} \times \mathbb{R}^n \), as claimed. This completes the proof of Theorem 2.4.

**Proof of Theorem 2.5**

The subsequent proofs will build on the following two lemmas.
Lemma A.3. For any $x \in E$ we have the identity

$$\text{span} \{\nabla F(\tau, x) : \tau \geq 0\} = \text{span} \left\{ (\kappa^\top)^p \psi - \frac{\psi^\top \kappa^p (x - \theta)}{\phi + \psi^\top x} \psi : p = 1, \ldots, d \right\}. $$

Proof. Expanding $F(\tau, x)$ as a power series in $\tau$ gives

$$F(\tau, x) = e^{-\alpha \tau} \left( 1 + \sum_{p \geq 1} \frac{\psi^\top \kappa^p (x - \theta) (-\tau)^p}{p!} \phi + \psi^\top x \right). $$

(29)

The lemma now follows from plugging (29) in expression (24) for $\nabla F(\tau, x)$. □

Lemma A.4. Assume $\phi + \psi^\top \theta \neq 0$, and consider any $x \in E$. The following conditions are equivalent:

(i) The vectors $(\kappa^\top)^p \psi, p = 1, \ldots, d,$ are linearly independent;

(ii) $\kappa$ is invertible and $\mathcal{U} = \{0\}$;

(iii) $\text{span} \{\nabla F(\tau, x) : \tau \geq 0\} = \mathbb{R}^d$.

Moreover, conditions (i) and (ii) are equivalent even if $\psi + \psi^\top \theta = 0$.

Proof. The equivalence of (i) and (ii) is deduced from the identity

$$\det (\kappa^\top \psi \cdots (\kappa^\top)^d \psi) = \det (\kappa^\top) \det (\psi \cdots (\kappa^\top)^{d-1} \psi),$$

together with Theorem 2.2, which in particular states that the second determinant on the right side is zero if and only if $\mathcal{U} = \{0\}$.

We next prove that (i) implies (iii). Since $(\kappa^\top)^p \psi, p = 1, \ldots, d,$ span $\mathbb{R}^d$, we can find $a_1, \ldots, a_d$ such that $\psi = \sum_{p=1}^d a_p (\kappa^\top)^p \psi$. Together with the representation in Lemma A.3 we deduce that the vector

$$\sum_{p=1}^d a_p \left[ (\kappa^\top)^p \psi - \frac{\psi^\top \kappa^p (x - \theta)}{\phi + \psi^\top x} \psi \right] = \psi - \frac{\psi^\top (x - \theta)}{\phi + \psi^\top x} \psi = \frac{\phi + \psi^\top \theta}{\phi + \psi^\top x} \psi$$

lies in $\text{span} \{\nabla F(\tau, x) : \tau \geq 0\}$. So does $\psi$ since $\phi + \psi^\top \theta \neq 0$, and it follows that we have $\text{span} \{\nabla F(\tau, x) : \tau \geq 0\} \supseteq \text{span} \{(\kappa^\top)^p \psi : p = 1, \ldots, d\} = \mathbb{R}^d$. This proves (iii).

It remains to prove that (iii) implies (i). To this end, we use the hypothesis together with the representation in Lemma A.3 to find $a_1, \ldots, a_p$ such that

$$\psi = \sum_{p=1}^d a_p \left[ (\kappa^\top)^p \psi - \frac{\psi^\top \kappa^p (x - \theta)}{\phi + \psi^\top x} \psi \right].$$
Re-arranging this expression and writing $u = \sum_{p=1}^{d} a_p (\kappa^\top)^p \psi$ yields

$$
\psi \left( 1 + \frac{u^\top (x - \theta)}{\phi + \psi^\top x} \right) = u.
$$

It follows that $\psi$ lies in $\text{span}\{(\kappa^\top)^p \psi : p = 1, \ldots, d\}$, which consequently is equal to $\text{span}\{\nabla F(\tau,x) : \tau \geq 0\} = \mathbb{R}^d$. This proves that (i) holds.

We can now proceed to the proof of Theorem 2.5. We first assume $\mathcal{U} = \{0\}$, so that $m = d$. Let $x, y \in E$. In view of the power series expansion (29) we have $F(\tau,x) = F(\tau,y)$ for all $\tau \geq 0$ if and only if

$$
\psi^\top (x - \theta) (\phi + \psi^\top x) = \psi^\top (y - \theta) (\phi + \psi^\top y), \quad p \geq 1.
$$

(30)

To prove sufficiency, assume $\kappa$ is invertible and $\phi + \psi^\top \theta \neq 0$. Then pick $x, y$ such that (30) is satisfied; we must prove that $x = y$. Lemma A.4 implies that the vectors $\kappa^\top \psi, \ldots, (\kappa^\top)^d \psi$ span $\mathbb{R}^d$, so we may find coefficients $a_1, \ldots, a_d$ so that $\psi = \sum_{p=1}^{d} a_p (\kappa^\top)^p \psi$. Multiplying both sides of (30) by $a_p$ and summing over $p = 1, \ldots, d$ yields

$$
\frac{\psi^\top (x - \theta)}{\phi + \psi^\top x} = \frac{\psi^\top (y - \theta)}{\phi + \psi^\top y},
$$

or, equivalently, $\psi^\top (x - y) (\phi + \psi^\top \theta) = 0$. Since $\phi + \psi^\top \theta \neq 0$ we then deduce from (30) that $\psi^\top (x - y) = 0$ for all $p \geq 0$, which by the aforementioned spanning property of $\psi^\top \kappa^p, p \geq 0$, implies $x = y$ as required. This finishes the proof of the first assertion.

To prove necessity, we argue by contradiction and suppose it is not true that $\kappa$ is invertible and $\phi + \psi^\top \theta \neq 0$. There are two cases. First, assume $\kappa$ is not invertible. We claim that there is an element $\eta \in \ker \kappa$ such that $\theta + s_1 \eta$ lies in the set $\{x \in \mathbb{R}^d : \phi + \psi^\top x \neq 0\}$ for all large $s$. Indeed, if this were not the case we would have $\ker \kappa \subseteq \ker \psi^\top$, which would contradict $\mathcal{U} = \{0\}$. So such an $\eta$ exists. Now simply take $x = \theta + s_1 \eta, y = \theta + s_2 \eta$ for large enough $s_1 \neq s_2$—clearly (30) holds for this choice, proving that injectivity fails.

The second case is where $\kappa$ is invertible, but $\phi + \psi^\top \theta = 0$. In particular $\phi + \psi^\top x = \psi^\top (x - \theta)$. Together with the fact that $\kappa^\top \psi, \ldots, (\kappa^\top)^d \psi$ span $\mathbb{R}^d$ (see Lemma A.4), this shows that (30) is equivalent to

$$
\frac{x - \theta}{\psi^\top (x - \theta)} = \frac{y - \theta}{\psi^\top (y - \theta)}.
$$

We deduce that $F(\tau,x)$ is constant along rays of the form $\theta + s(x - \theta)$, where $x$ is any point in the state space, and thus that injectivity fails.
The proof of the theorem is now complete for the case $U = \{0\}$. The general case where $U$ is not necessarily trivial follows by applying the $U = \{0\}$ case to the model with factor process $Z$ and state price density $\zeta_t = e^{-\alpha t}(\hat{\phi} + \hat{\psi}^\top Z_t)$. Indeed, this model has a trivial term structure kernel in view of Theorem 2.4 (iii).

**Proof of Theorem 2.6**

The proof uses the following identity from Fourier analysis, valid for any $\mu > 0$ and $s \in \mathbb{R}$ (see for instance Bateman and Erdélyi (1954, Formula 3.2(3))):

$$s^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(\mu + i\lambda)s} \frac{1}{(\mu + i\lambda)^2} d\lambda. \quad (31)$$

Let $q(ds)$ denote the conditional distribution of the random variable $p_{swap}(X_{T_0})$, given $\mathcal{F}_t$, so that

$$\hat{q}(z) = \int_{\mathbb{R}} e^{zs} q(ds)$$

for every $z \in \mathbb{C}$ such that the right side is well-defined and finite. Pick $\mu > 0$ such that $\int_{\mathbb{R}} e^{\mu}s q(ds) < \infty$. Then,

$$\int_{\mathbb{R}^2} e^{(\mu + i\lambda)s} \frac{1}{(\mu + i\lambda)^2} d\lambda \otimes q(ds) = \int_{\mathbb{R}^2} \frac{e^{\mu}s}{\mu^2 + \lambda^2} d\lambda \otimes q(ds) = \int_{\mathbb{R}} e^{\mu}s q(ds) \int_{\mathbb{R}} \frac{1}{\mu^2 + \lambda^2} d\lambda < \infty,$$

where the second equality follows from Tonelli’s theorem. This justifies applying Fubini’s theorem in the following calculation, which uses the identity (31) on the second line:

$$\mathbb{E} [p_{swap}(X_{T_0})^+ | \mathcal{F}_t] = \int_{\mathbb{R}} s^+ q(ds)$$

$$= \int_{\mathbb{R}} \left( \frac{1}{2\pi} \int_{\mathbb{R}} e^{(\mu + i\lambda)s} \frac{1}{(\mu + i\lambda)^2} d\lambda \right) q(ds)$$

$$= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} d\lambda$$

$$= \frac{1}{\pi} \int_{0}^{\infty} \text{Re} \left[ \frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda.$$
Proof of Theorem 3.2

The proof of Theorem 3.2 requires some notation and an additional lemma. For a multiindex \( \mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}_0^d \) we write \( |\mathbf{k}| = k_1 + \cdots + k_d \), \( x^\mathbf{k} = x_1^{k_1} \cdots x_d^{k_d} \), and \( \partial^\mathbf{k} = \partial^{|\mathbf{k}|}/\partial x_1^{k_1} \cdots \partial x_d^{k_d} \).

Lemma A.5. Assume \( \sigma(X_t) \) is invertible \( dt \otimes d\mathbb{P} \)-almost surely. For any function \( f \in C^{1,\infty}(\mathbb{R}_+ \times E) \) we have

\[
\{ f(t, X_t) = 0 \} \subseteq \bigcap_{\mathbf{k} \in \mathbb{N}_0^d} \{ \partial^\mathbf{k} f(t, X_t) = 0 \},
\]

up to a \( dt \otimes d\mathbb{P} \)-nullset.

Proof. Let \( n \geq 0 \) and suppose we have, up to a nullset,

\[
\{ f(t, X_t) = 0 \} \subseteq \{ \partial^\mathbf{k} f(t, X_t) = 0 \}
\]

for all \( \mathbf{k} \in \mathbb{N}_0^d \) with \( |\mathbf{k}| = n \). Fix such a \( \mathbf{k} \) and set \( g(t, x) = \partial^\mathbf{k} f(t, x) \). The occupation time formula, see Revuz and Yor (1999, Corollary VI.1.6), yields

\[
1_{\{g(t,X_t)=0\}} \nabla g(t,X_t)^\top a(X_t) \nabla g(t,X_t) = 0 \quad dt \otimes d\mathbb{P}\text{-a.s.}
\]

This implies \( \{ g(t, X_t) = 0 \} \subseteq \{ \nabla g(t, X_t)^\top a(X_t) \nabla g(t, X_t) = 0 \} \) up to a nullset. Since \( a(X_t) \) is invertible \( dt \otimes d\mathbb{P} \)-almost surely we get, again up to a nullset,

\[
g(t, X_t) = 0 \subseteq \{ \nabla g(t, X_t) = 0 \}.
\]

We deduce that (32) holds for all \( \mathbf{k} \in \mathbb{N}_0^d \) with \( |\mathbf{k}| = n + 1 \). Since (32) is trivially true for \( n = 0 \), the result follows by induction. \( \Box \)

We can now prove Theorem 3.2. By a standard argument involving the martingale representation theorem, bond market completeness holds if and only if for any given \( T \geq 0 \) there exist maturities \( T_i \geq T, i = 1, \ldots, m \), such that \( dt \otimes d\mathbb{P} \)-almost surely the volatility vectors \( \nu(T_i - t, X_t) \) span \( \mathbb{R}^d \). Since \( \sigma(X_t) \) is invertible \( dt \otimes d\mathbb{P} \)-almost surely this happens if and only if \( dt \otimes d\mathbb{P} \)-almost surely the vectors \( \nabla F(T_i - t, X_t) \) span \( \mathbb{R}^d \). This shows, in particular, that (i) implies (ii).

To prove that (ii) implies (iii), suppose (iii) fails. Lemma A.3 then implies that for each \( x \in E \), span\{\( \nabla F(\tau, x) : \tau \geq 0 \)\} is not all of \( \mathbb{R}^d \). Thus (ii) fails.

It remains to prove that (iii) implies (i), so we assume that \( \kappa \) is invertible and \( \mathcal{U} = \{0\} \). Choose maturities \( T_1, \ldots, T_d \) greater than or equal to \( T \) so that the function

\[
g(t, x) = \det \left( \nabla F(T_i - t, x) \cdots \nabla F(T_d - t, x) \right)
\]
is not identically zero. This is possible by Lemma A.3. A calculation yields
\[ \nabla F(\tau, x) = e^{-\alpha \tau} \left( \psi + \psi^\top x \right) \eta(\tau, x), \]
where \( \eta(\tau, x) \) is a vector of first degree polynomials in \( x \) whose coefficients are analytic functions of \( \tau \). Defining
\[ f(t, x) = \det \left( \eta(T_1 - t, x) \cdots \eta(T_d - t, x) \right), \]
we have \( g(t, x) = 0 \) if and only if \( f(t, x) = 0 \). Hence \( f(t, x) \) is not identically zero. Our goal is to strengthen this to the statement that \( \{ f(t, X_t) = 0 \} \) is a \( dt \otimes d\mathbb{P} \)-nullset. Indeed, then \( \{ g(t, X_t) = 0 \} \) is also a \( dt \otimes d\mathbb{P} \)-nullset, implying that completeness holds. To prove that \( \{ f(t, X_t) = 0 \} \) is a \( dt \otimes d\mathbb{P} \)-nullset, note that \( f(t, x) \) is of the form
\[ f(t, x) = \sum_{|k| \leq n} c_k(t) x^k, \]
where \( n = \max_{0 \leq t \leq T} \deg f(t, \cdot) < \infty \). Lemma A.5 implies
\[ \{ f(t, X_t) = 0 \} \subseteq \bigcap_{|k| = n} \{ c_k(t) = 0 \}, \]
up to a nullset. Assume for contradiction that the left side is of positive \( dt \otimes d\mathbb{P} \)-measure. Then so is the right side, whence all the \( c_k \) (which are deterministic) vanish on a \( t \)-set of positive Lebesgue measure. The zero set of each \( c_k \) must thus contain an accumulation point, so that, by analyticity, they are all identically zero, see Rudin (1987, Theorem 10.18). Hence we have either \( \max_{0 \leq t \leq T} \deg f(t, \cdot) \leq n - 1 \) (if \( n \geq 1 \)) or \( f(t, x) \equiv 0 \) (if \( n = 0 \)). In both cases we obtain a contradiction, which shows that \( \{ f(t, X_t) = 0 \} \) is a \( dt \otimes d\mathbb{P} \)-nullset, as required.

Finally, the equivalence of (iii) and (iv) follows easily from Theorem 2.5. The theorem is proved.

**Proof of Corollary 3.4**

Theorem A.7 below gives a description of the variance-covariance kernel \( \mathcal{W} \). Its proof requires the following lemma.

**Lemma A.6.** Assume \( \phi + \psi^\top \theta \neq 0 \), and consider any \( x \in E \). The following conditions are equivalent.
(i) $\psi \in \text{span}\{(\kappa^\top)p\psi : p = 1, \ldots, d\}$,
(ii) $U^\perp = \text{span}\{\nabla F(\tau, x) : \tau \geq 0\}$.

Proof. The proof is a straightforward adaptation of the proof of the equivalence of (i) and (iii) in Lemma A.4, and therefore omitted.

**Theorem A.7.** The variance-covariance kernel satisfies

$$U \cap \mathcal{W} \subseteq U \cap \bigcap_{\eta \in U^\perp} \ker \eta^\top a(\cdot) \eta$$

with equality if $\phi + \psi^\top \theta \neq 0$ and $\psi \in \text{span}\{(\kappa^\top)p\psi : p = 1, \ldots, d\}$.

Proof. Consider an arbitrary vector $\xi \in U \cap \mathcal{W}$. Since $F(\tau, x + s\xi)$ is constant in $s$, we have that $\nabla G(\tau_1, \tau_2, x)^\top \xi = 0$ if and only if $\nabla \overline{G}(\tau_1, \tau_2, x)^\top \xi = 0$, where we define

$$\overline{G}(\tau_1, \tau_2, x) = \nabla F(\tau_1, x)^\top a(x) \nabla F(\tau_2, x).$$

Now, for any $x \in E$ and $\tau \geq 0$, the chain rule yields

$$\left.\frac{d}{ds} \nabla F(\tau, x + s\xi)\right|_{s=0} = \nabla F(\tau, x)^\top = 0.$$

Hence, by the product rule,

$$\left.\frac{d}{ds} \overline{G}(\tau_1, \tau_2, x + s\xi)\right|_{s=0} = \left.\frac{d}{ds} \nabla F(\tau_1, x)^\top a(x + s\xi) \nabla F(\tau_2, x)\right|_{s=0}.$$

The right side is zero for all $x \in E$, $\tau_1, \tau_2 \geq 0$ if and only if, for every $x \in E$,

$$\left.\frac{d}{ds} \eta_1^\top a(x + s\xi) \eta_2\right|_{s=0} = 0$$

holds for all $\eta_1, \eta_2 \in \text{span}\{\nabla F(\tau, x) : \tau \geq 0\}$. But we always have

$$\text{span}\{\nabla F(\tau, x) : \tau \geq 0\} \subseteq U^\perp,$$

with equality if $\phi + \psi^\top \theta \neq 0$ and $\psi \in \text{span}\{(\kappa^\top)p\psi : p = 1, \ldots, d\}$, due to Lemma A.6. This proves that the inclusion

$$\mathcal{W} \subseteq \left\{ \xi \in U : \left.\frac{d}{ds} \eta_1^\top a(x + s\xi) \eta_2\right|_{s=0} = 0 \text{ for all } x \in E, \eta_1, \eta_2 \in U^\perp \right\}$$

(33)
holds, with equality if $\phi + \psi^\top \theta \neq 0$ and $\psi \in \text{span}\{(\kappa^\top)p^\top : p = 1, \ldots, d\}$. Finally, the identity
\[
\eta_1^\top A \eta_2 = \frac{1}{2}\left[\eta_1^\top A \eta_1 + \eta_2^\top A \eta_2 - (\eta_1 - \eta_2)^\top A(\eta_1 - \eta_2)\right],
\]
valid for any symmetric matrix $A$, implies that the right side of (33) is equal to
\[
\left\{\xi \in \mathcal{U} : \left.\frac{d}{ds} \eta^\top a(x + s\xi)\eta\right|_{s=0} = 0 \text{ for all } x \in E, \eta \in \mathcal{U}^\perp\right\}.
\]
Since this set is equal to $\mathcal{U} \cap \bigcap_{\eta_1, \eta_2 \in \mathcal{U}^\perp} \ker \eta_1^\top a(\cdot) \eta_2 \cap \mathcal{U}$, the proof of Theorem A.7 is complete.

**Remark A.8.** Note that the set on the right side in Theorem A.7 is equal to
\[
\bigcap_{\eta_1, \eta_2 \in \mathcal{U}^\perp} \ker \eta_1^\top a(\cdot) \eta_2 \cap \mathcal{U},
\]
see (33).

Building on this remark we now discuss how the USV factors and residual factors affect the volatility of the transformed factor process $\hat{X}_t = (Z_t, U_t)$, where $U_t = (V_t, W_t)$. The dynamics of $\hat{X}_t$ can be written
\[
\frac{d}{dt} \begin{pmatrix} Z_t \\ U_t \end{pmatrix} = \begin{pmatrix} \hat{\kappa}_{ZZ}(\hat{\theta}_Z - Z_t) \\ \hat{\kappa}_{ZU}(\hat{\theta}_Z - Z_t) + \hat{\kappa}_{UU}(\hat{\theta}_U - U_t) \end{pmatrix} dt + \hat{\sigma}(Z_t, V_t, W_t)dB_t,
\]
where $\hat{\kappa}$ and $\hat{\theta}$ are given by (9), and $\hat{\sigma}(z, v, w) = S\sigma(S^{-1}(z, v, w))$. The corresponding diffusion matrix is
\[
\hat{a}(z, v, w) = \hat{\sigma}\hat{\sigma}^\top(z, v, w) = Sa(S^{-1}(z, v, w))S^\top.
\]
Consider now those components of $\hat{a}(z, v, w)$ that are related to the volatility and covariation of the term structure factors $Z_t$,
\[
\hat{a}_{ij}(z, v, w) = e_i^\top \hat{a}(z, v, w)e_j = \eta_i^\top a(S^{-1}\hat{x})\eta_j, \quad i, j \in \{1, \ldots, m\},
\]
where we set $\eta_i = S^\top e_i \in \mathcal{U}^\perp$. In view of Theorem A.7 and Remark A.8, we see that in the generic case, the functions $\hat{a}_{ij}(z, v, w)$ are all constant in $w$, but not all constant in $v$. In other words, the volatilities and covariations of the term structure factors $Z_t$ are not directly affected by the residual factors $W_t$, but are affected by the USV factors $V_t$. This fact provides a further justification for our terminology, and it proves Corollary 3.4 in particular.
Proof of Theorem 4.1

We first prove a slightly weaker result, which is valid for linear-rational term structure models (2)–(3) with a semimartingale factor process $X_t$ whose minimal state space is the nonnegative orthant $\mathbb{R}_+^d$. Hereby, we say that the state space $E$ is minimal if $\mathbb{P}(X_t \in U$ for some $t \geq 0) > 0$ holds for any relatively open subset $U \subset E$.

**Lemma A.9.** Assume $X_t$ is a semimartingale of the form (2) whose minimal state space is $\mathbb{R}_+^d$. Then $\kappa_{ij} \leq 0$ for all $i \neq j$.

**Proof.** Let $G(\tau, x)$ denote the solution to the linear differential equation

$$\partial_\tau G(\tau, x) = \kappa(\theta - G(\tau, x)), \quad G(0, x) = x,$$

so that, by Lemma A.1, $\mathbb{E}[X_{\rho+\tau} \mid \mathcal{F}_\rho] = G(\tau, X_{\rho})$ holds for any bounded stopping time $\rho$ and any (deterministic) $\tau \geq 0$. Pick $i, j \in \{1, \ldots, d\}$ with $i \neq j$, and assume for contradiction that $\kappa_{ij} > 0$. Then, for $\lambda > 0$ large enough, we have

$$\partial_\tau G_i(0, \lambda e_j) = e_i^\top \kappa(\theta - \lambda e_j) = e_i^\top \kappa \theta - \lambda \kappa_{ij} < -2,$$

where $e_i$ ($e_j$) denotes the $i$:th ($j$:th) unit vector, and $G_i$ is the $i$:th component of $G$. By continuity there is some $\varepsilon > 0$ such that $\partial_\tau G_i(\tau, \lambda e_j) \leq -2$ for all $\tau \in [0, 2\varepsilon]$.

Hence $G_i(\tau, \lambda e_j) = 0 + \int_0^\tau \partial_\tau G_i(s, \lambda e_j)ds \leq -2\tau$ for all $\tau \in [0, 2\varepsilon]$. By continuity of $(\tau, x) \mapsto G(\tau, x)$ there is some $r > 0$ such that

$$G_i(\tau, x) \leq -\tau \quad \text{holds for all } \tau \in [0, \varepsilon], \ x \in B(\lambda e_j, r),$$

where $B(x, r)$ is the ball of radius $r$ centered at $x$. Now define

$$\rho = n \wedge \inf\{t \geq 0 : X_t \in B(\lambda e_j, r)\}, \quad A = \{X_{\rho} \in B(\lambda e_j, r)\},$$

where $n$ is chosen large enough that $\mathbb{P}(A) > 0$. The assumption that $\mathbb{R}_+^d$ is a minimal state space implies that such an $n$ exists. Then

$$\mathbb{E}[1_A X_{\iota, \rho+\varepsilon}] = \mathbb{E}[1_A \mathbb{E}[X_{\iota, \rho+\varepsilon} \mid \mathcal{F}_\rho]] = \mathbb{E}[1_A G_i(\varepsilon, X_{\rho})] \leq -\varepsilon \mathbb{P}(A) < 0,$$

whence $\mathbb{P}(X_{\iota, \rho+\varepsilon} < 0) > 0$, which is the desired contradiction. The lemma is proved. \hfill $\square$

Now consider a linear-rational term structure model (2)–(3) with a semimartingale factor process $X_t$ and minimal state space $E = \mathbb{R}_+^d$. Since $\phi + \psi^\top x$ is assumed positive on $\mathbb{R}_+^d$, we must have $\psi \in \mathbb{R}_+^d$ and $\phi > 0$. Dividing $\zeta_t$ by $\phi$ does not affect
any model prices, so we may take $\phi = 1$. Moreover, after permuting and scaling the components, $X_1$ is still a semimartingale of the form (2) with minimal state space $\mathbb{R}^d_+$, so we can assume $\psi = 1_p$. Here, we let $1_p$ denote the vector in $\mathbb{R}^d$ whose first $p$ ($p \leq d$) components are ones, and the remaining components are zeros. As before, we write $1 = 1_d$, whenever there is no ambiguity. The short rate is then given by

$$r_t = \alpha - \rho(X_t),$$

where

$$\rho(x) = \frac{1_p^\top \kappa \theta - 1_p^\top \kappa x}{1 + 1_p^\top x} = \frac{1_p^\top \kappa \theta + \sum_{i=1}^{d} (-1_p^\top \kappa_i x_i)}{1 + \sum_{i=1}^{p} x_i}$$

and where $\kappa_i$ denotes the $i$th column vector of $\kappa$.

**Lemma A.10.** The short rates are bounded from below, $\alpha^* = \sup_{x \in \mathbb{R}^d_+} \rho(x) < \infty$, if and only if $1_p^\top \kappa_i = 0$ for $i > p$. In this case, $\alpha^* = \max S_p$ and $\alpha_* = \min S_p$ where

$$S_p = \{1_p^\top \kappa \theta, -1_p^\top \kappa_1, \ldots, -1_p^\top \kappa_p\}.$$

and the submatrix $\kappa_{1\ldots p, p+1\ldots d}$ is zero.

**Proof.** Lemma A.9 implies that $-1_p^\top \kappa_i \geq 0$ for $i > p$. Since the state components $x_i$ are nonnegative and unbounded from above, it is then obvious that $\alpha^*$ is finite if and only if $1_p^\top \kappa_i = 0$ for $i > p$. This means that the submatrix $\kappa_{1\ldots p, p+1\ldots d}$ is zero.

The expressions for $\alpha^*$ and $\alpha_*$ now follow by observing that for each $x \in \mathbb{R}^d_+$, $\rho(x)$ is a convex combination of the elements in $S_p$. \qed

We can now prove Theorem 4.1. To this end, we first observe that the factor process $X_t$ remains a square-root process after coordinatewise scaling and permutation of its components. Hence, as above, we can assume that $\phi = 1$ and $\psi = 1_p$ for some $p \leq d$. Lemma A.10 then shows that short rates are bounded from below if and only if submatrix $\kappa_{1\ldots p, p+1\ldots d}$ vanishes. If $p = d$ there is nothing left to prove. So assume now that $p < d$. This implies that $(X_{1t}, \ldots, X_{pt})$ is an autonomous square-root process on the smaller state space $\mathbb{R}^p_d$. Since the state price density $\zeta_t$ only depends on the first $p$ components of $X_t$, the pricing model is unaffected if we exclude the last $d - p$ components, and this proves that we may always take $p = d$, as desired.

Finally, the expressions for $\alpha^*$ and $\alpha_*$ follow directly from Lemma A.10. This completes the proof of Theorem 4.1.

**Proof of Corollary 4.2**

The assertion about $\dim \mathcal{U}$ follows from Theorem 2.4 and Lemma A.2. To prove that all unspanned factors are USV factors we apply Corollary 3.4. To this end it
suffices to consider the entries $\hat{a}_{ii}(z,u)$ of the diffusion matrix of the transformed factor process $\hat{X}_t = (Z_t, U_t)$. Using that $\hat{a}(z,u) = S\sigma\sigma^T(S^{-1}(z,u))S^T$, a calculation yields

$$\hat{a}_{ii}(z,u) = \sigma_i^2 z_i + (\sigma_{m+i}^2 - \sigma_i^2)u_i, \quad i = 1, \ldots, n,$$

which is non-constant in $u_i$ since $\sigma_i \neq \sigma_{m+i}$. Since this holds for all $i = 1, \ldots, n$, Corollary 3.4 implies that all unspanned factors are USV factors.

### B An Example with Residual Factors

We provide an example illustrating that residual factors may affect the distribution of future bond prices, despite having no instantaneous impact on the current term structure or bond return volatilities.

Consider a three-factor linear-rational model with state space $E = \mathbb{R}_+ \times \mathbb{R}^2$ and factor process given by

\[
\begin{align*}
    dX_{1t} &= \lambda_1(1 - X_{1t})dt + X_{1t}\varphi(X_{2t})dB_{1t} \\
    dX_{2t} &= -\lambda_2X_{2t}dt + X_{3t}dB_{2t} \\
    dX_{3t} &= -\lambda_3X_{3t}dt + dB_{3t},
\end{align*}
\]

where $\lambda_i > 0$ ($i = 1, 2, 3$), and $\varphi : \mathbb{R} \to [1, 2]$ is a strictly increasing, differentiable function with $s\varphi'(s)$ bounded. Then $x \mapsto x_1\varphi(x_2)$ is Lipschitz continuous, and it follows that there is a unique strong solution to the above equation, starting from any point $x \in E$. Note that $X_{1t}$ necessarily stays nonnegative since its drift is positive and the diffusion component vanishes at the origin. The state price density is taken to be

$$\zeta_t = e^{-\alpha t}(1 + X_{1t})$$

for some $\alpha$. The short rate is then given by $r_t = \alpha - \lambda_1(1 - X_{1t})/(1 + X_{1t})$, see (6), and therefore, see (7), we pick

$$\alpha = \alpha^* = \sup_{x \in E} \lambda_1 \frac{1 - x_1}{1 + x_1} = \lambda_1.$$

Since also $\alpha^* = -\lambda_1$, this gives a short rate contained in $[0, 2\lambda_1]$. What are the unspanned factors in this model? The term structure kernel, given by (8), reduces to

$$U = \{\xi \in \mathbb{R}^3 : \xi_1 = 0\} = \text{span} \{e_2, e_3\}.$$

Furthermore, the transformation $S$ can be taken to be the identity. Thus $X_{2t}$ is a USV factor by Corollary 3.4, since $a_{11}(x + se_2) = x_1^2\varphi(x_2 + s)^2$ is non-constant in $s$. On
the other hand, $X_{3t}$ is a residual factor, since $a_{11}(x + se_3) = x_1^2\varphi(x_2)^2$ is constant in $s$. At the same time, however, $a_{22}(x + se_3)$ varies with $s$. The message is the following: while a perturbation of the residual factor $X_{3t}$ has no immediate impact on the bond prices or their volatilities and covariances, such a perturbation does affect the volatility of the USV factor $X_{2t}$. Therefore it affects the future distribution of this factor, and hence also the future distribution of the volatility of the term structure factor $X_{1t}$. This in turn alters the future distribution of bond prices. In conclusion, derivatives prices may, in general, be sensitive to residual factors, and thus contain information about their current values. This may happen despite the fact that the residual factors are neither term structure factors, nor USV factors.

C Quasi-Maximum Likelihood Estimation

We estimate by quasi-maximum likelihood in conjunction with Kalman filtering. For this purpose, we cast the model in state space form with a measurement equation describing the relation between the state variables and the observable swap rates and swaption IVs, as well as a transition equation describing the discrete-time dynamics of the state variables.

Let $X_t$ denote the vector of state variables and let $Y_t$ denote the vector consisting of the term structure of swap rates and swaption IVs observed at time $t$. The measurement equation is given by

$$Y_t = h(X_t; \Theta) + u_t, \quad u_t \sim N(0, \Sigma), \quad (35)$$

where $h$ is the pricing function, $\Theta$ is the vector of model parameters, and $u_t$ is a vector of i.i.d. Gaussian pricing errors with covariance matrix $\Sigma$. To reduce the number of parameters in $\Sigma$, we assume that the pricing errors are cross-sectionally uncorrelated (that is, $\Sigma$ is diagonal), and that one variance, $\sigma_{rates}^2$, applies to all pricing errors for swap rates, and that another variance, $\sigma_{swaption}^2$, applies to all pricing errors for swaption IVs.

While the transition density of $X_t$ is unknown, its conditional mean and variance is known in closed form, because $X_t$ is a square-root process. We approximate the transition density with a Gaussian density with identical first and second moments, in which case the transition equation is of the form

$$X_t = \Phi_0 + \Phi_X X_{t-1} + w_t, \quad w_t \sim N(0, Q_t), \quad (36)$$

where $Q_t$ is a linear function of $X_{t-1}$. 

45
As both swap rates and swaption IVs are non-linearly related to the state variables, we apply the nonlinear unscented Kalman filter.\textsuperscript{24} The Kalman filter produces one-step-ahead forecasts for $Y_t$, $\hat{Y}_{t|t-1}$, and the corresponding error covariance matrices, $F_{t|t-1}$, from which we construct the log-likelihood function

$$L(\Theta) = -\frac{1}{2} \sum_{t=1}^{T} \left( n_t \log 2\pi + \log |F_{t|t-1}| + (Y_t - \hat{Y}_{t|t-1})^T F_{t|t-1}^{-1} (Y_t - \hat{Y}_{t|t-1}) \right), \quad (37)$$

where $T$ is the number of observation dates and $n_t$ is the number of observations in $Y_t$. The (quasi) maximum likelihood estimator, $\hat{\Theta}$, is then

$$\hat{\Theta} = \arg \max_{\Theta} L(\Theta). \quad (38)$$

\textsuperscript{24}Leippold and Wu (2007) appear to be the first to apply the unscented Kalman filter to the estimation of dynamic term structure models. Christoffersen, Jacobs, Karoui, and Mimouni (2009) show that it has very good finite-sample properties when estimating models using swap rates.
Table 1: Summary statistics.
The table reports the mean, median, standard deviation, minimum, and maximum of each time series. Swap rates are reported in percentages. Swaption normal implied volatilities are reported in basis points. Each time series consists of 866 weekly observations from January 29, 1997 to August 28, 2013.
<table>
<thead>
<tr>
<th></th>
<th>$LRSQ(3,1)$</th>
<th>$LRSQ(3,2)$</th>
<th>$LRSQ(3,3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_{1,1}$</td>
<td>0.0640</td>
<td>0.2286</td>
<td>0.2151</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0023)</td>
<td>(0.0015)</td>
</tr>
<tr>
<td>$\kappa_{2,2}$</td>
<td>0.4389</td>
<td>0.1335</td>
<td>0.1863</td>
</tr>
<tr>
<td></td>
<td>(0.0050)</td>
<td>(0.0010)</td>
<td>(0.0013)</td>
</tr>
<tr>
<td>$\kappa_{3,3}$</td>
<td>0.1611</td>
<td>1.1948</td>
<td>0.6931</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
<td>(0.0122)</td>
<td>(0.0040)</td>
</tr>
<tr>
<td>$\kappa_{2,1}$</td>
<td>$-0.1268$</td>
<td>$-0.2867$</td>
<td>$-0.2688$</td>
</tr>
<tr>
<td></td>
<td>(0.0010)</td>
<td>(0.0025)</td>
<td>(0.0016)</td>
</tr>
<tr>
<td>$\kappa_{3,2}$</td>
<td>$-0.5016$</td>
<td>$-0.1916$</td>
<td>$-0.2400$</td>
</tr>
<tr>
<td></td>
<td>(0.0052)</td>
<td>(0.0011)</td>
<td>(0.0012)</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.8333</td>
<td>0.1294</td>
<td>0.1688</td>
</tr>
<tr>
<td></td>
<td>(0.0349)</td>
<td>(0.0097)</td>
<td>(0.0037)</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.2827</td>
<td>0.2778</td>
<td>0.2436</td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
<td>(0.0083)</td>
<td>(0.0028)</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>0.8802</td>
<td>0.0876</td>
<td>0.0844</td>
</tr>
<tr>
<td></td>
<td>(0.0214)</td>
<td>(0.0016)</td>
<td>(0.0011)</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>0.1458</td>
<td>0.1249</td>
<td>0.0809</td>
</tr>
<tr>
<td></td>
<td>(0.0271)</td>
<td>(0.0056)</td>
<td>(0.0014)</td>
</tr>
<tr>
<td>$\theta_5$</td>
<td>0.2681</td>
<td>0.1167</td>
<td>0.1167</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0015)</td>
<td></td>
</tr>
<tr>
<td>$\theta_6$</td>
<td></td>
<td></td>
<td>0.0404</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0007)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.2326</td>
<td>0.2366</td>
<td>0.2347</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0058)</td>
<td>(0.0042)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.7095</td>
<td>0.2297</td>
<td>0.2478</td>
</tr>
<tr>
<td></td>
<td>(0.0047)</td>
<td>(0.0015)</td>
<td>(0.0013)</td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td>0.1244</td>
<td>0.0514</td>
<td>0.0507</td>
</tr>
<tr>
<td></td>
<td>(0.0016)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>$\sigma_4$</td>
<td>1.8084</td>
<td>0.9593</td>
<td>0.9486</td>
</tr>
<tr>
<td></td>
<td>(0.0153)</td>
<td>(0.0098)</td>
<td>(0.0089)</td>
</tr>
<tr>
<td>$\sigma_5$</td>
<td>0.7495</td>
<td>0.7801</td>
<td>0.7801</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0057)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_6$</td>
<td></td>
<td></td>
<td>0.2880</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0020)</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>$-0.0007$</td>
<td>0.7767</td>
<td>0.6337</td>
</tr>
<tr>
<td></td>
<td>(0.2267)</td>
<td>(0.5230)</td>
<td>(0.0366)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>0.0403</td>
<td>$-2.8330$</td>
<td>$-1.6450$</td>
</tr>
<tr>
<td></td>
<td>(0.2971)</td>
<td>(0.4666)</td>
<td>(0.1438)</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>$-2.0008$</td>
<td>19.0997</td>
<td>8.7441</td>
</tr>
<tr>
<td></td>
<td>(0.3637)</td>
<td>(1.2974)</td>
<td>(0.7070)</td>
</tr>
<tr>
<td>$\delta_4$</td>
<td>0.0140</td>
<td>0.1103</td>
<td>0.1641</td>
</tr>
<tr>
<td></td>
<td>(0.2821)</td>
<td>(0.5150)</td>
<td>(0.0218)</td>
</tr>
<tr>
<td>$\delta_5$</td>
<td>$-1.3079$</td>
<td>$-1.5946$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.6096)</td>
<td>(0.0380)</td>
<td></td>
</tr>
<tr>
<td>$\delta_6$</td>
<td></td>
<td>$-1.8039$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.1014)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Maximum likelihood estimates.
Table 2: Maximum likelihood estimates (cont.)

The table reports parameter estimates with asymptotic standard errors are in parentheses. \( \sigma_{rates} \) denotes the standard deviation of swap rate pricing errors and \( \sigma_{swaptions} \) denotes the standard deviation of swaption pricing errors in terms of normal implied volatilities. Both \( \sigma_{rates} \) and \( \sigma_{options} \) are measured in basis points. \( \alpha \) is chosen as the smallest value that guarantees a nonnegative short rate. \( supr_t \) is the upper bound on possible short rates. \( \mathcal{L} \) denotes the log-likelihood value. The sample period consists of 866 weekly observations from January 29, 1997 to August 28, 2013.

<table>
<thead>
<tr>
<th></th>
<th>( LRSQ(3,1) )</th>
<th>( LRSQ(3,2) )</th>
<th>( LRSQ(3,3) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{rates} )</td>
<td>7.2649</td>
<td>4.4289</td>
<td>4.2451</td>
</tr>
<tr>
<td></td>
<td>(0.0531)</td>
<td>(0.0441)</td>
<td>(0.0421)</td>
</tr>
<tr>
<td>( \sigma_{swaptions} )</td>
<td>8.2822</td>
<td>7.5703</td>
<td>5.1563</td>
</tr>
<tr>
<td></td>
<td>(0.0563)</td>
<td>(0.0494)</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.0627</td>
<td>0.0581</td>
<td>0.0537</td>
</tr>
<tr>
<td>( supr_t )</td>
<td>0.2238</td>
<td>1.2529</td>
<td>0.7468</td>
</tr>
<tr>
<td>( \mathcal{L} \times 10^{-4} )</td>
<td>5.6740</td>
<td>5.8779</td>
<td>6.0263</td>
</tr>
<tr>
<td>Specification</td>
<td>Full sample</td>
<td>Pre-ZLB</td>
<td>ZLB</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Swaps</td>
<td>Swaptions</td>
<td>Swaps</td>
</tr>
<tr>
<td>( LRSQ(3,1) )</td>
<td>5.71</td>
<td>7.21</td>
<td>5.16</td>
</tr>
<tr>
<td>( LRSQ(3,2) )</td>
<td>3.21</td>
<td>6.13</td>
<td>2.36</td>
</tr>
<tr>
<td>( LRSQ(3,3) )</td>
<td>3.29</td>
<td>4.20</td>
<td>2.65</td>
</tr>
<tr>
<td>( LRSQ(3,2)-LRSQ(3,1) )</td>
<td>(-2.50^{* * *})</td>
<td>(-1.07^{*})</td>
<td>(-2.80^{* * *})</td>
</tr>
<tr>
<td></td>
<td>(-6.33)</td>
<td>(-1.92)</td>
<td>(-5.65)</td>
</tr>
<tr>
<td>( LRSQ(3,3)-LRSQ(3,2) )</td>
<td>0.08</td>
<td>(-1.93^{* * *})</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(-5.90)</td>
<td>(3.22)</td>
</tr>
</tbody>
</table>

Table 3: Comparison of model specifications.
The table reports means of time series of the root mean squared pricing errors (RMSE) of swap rates and normal implied swaption volatilities. Units are basis points. \( t \)-statistics, corrected for heteroscedasticity and serial correlation up to 50 lags using the method of Newey and West (1987), are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. The full sample period consists of 866 weekly observations from January 29, 1997 to August 28, 2013. The zero lower bound (ZLB) sample period consists of 246 weekly observations after December 16, 2008. The pre-ZLB sample period consists of 620 weekly observations before December 16, 2008.
Table 4: Unconditional excess swap returns.
The table reports the annualized mean and volatility of nonoverlapping monthly excess returns on interest rate swaps with 1-month forward start. Also reported is the annualized Sharpe ratio (SR). Excess returns are in percent. The top panel shows results in the data, where each time series consists of 200 monthly observations from February 1997 to August 2013. The lower panels show results in simulated data, where each time series consists of 600,000 monthly observations (50,000 years).

<table>
<thead>
<tr>
<th></th>
<th>1 yr</th>
<th>2 yrs</th>
<th>3 yrs</th>
<th>5 yrs</th>
<th>7 yrs</th>
<th>10 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.97</td>
<td>1.79</td>
<td>2.44</td>
<td>3.36</td>
<td>3.89</td>
<td>4.31</td>
</tr>
<tr>
<td>Vol</td>
<td>0.80</td>
<td>1.76</td>
<td>2.76</td>
<td>4.58</td>
<td>6.09</td>
<td>8.00</td>
</tr>
<tr>
<td>SR</td>
<td>1.21</td>
<td>1.01</td>
<td>0.88</td>
<td>0.73</td>
<td>0.64</td>
<td>0.54</td>
</tr>
<tr>
<td><strong>LRSQ(3,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.42</td>
<td>0.81</td>
<td>1.15</td>
<td>1.72</td>
<td>2.13</td>
<td>2.55</td>
</tr>
<tr>
<td>Vol</td>
<td>0.59</td>
<td>1.32</td>
<td>2.20</td>
<td>3.92</td>
<td>5.30</td>
<td>6.67</td>
</tr>
<tr>
<td>SR</td>
<td>0.71</td>
<td>0.61</td>
<td>0.52</td>
<td>0.44</td>
<td>0.40</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>LRSQ(3,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.41</td>
<td>0.84</td>
<td>1.21</td>
<td>1.76</td>
<td>2.14</td>
<td>2.47</td>
</tr>
<tr>
<td>Vol</td>
<td>0.64</td>
<td>1.40</td>
<td>2.21</td>
<td>3.69</td>
<td>4.96</td>
<td>6.44</td>
</tr>
<tr>
<td>SR</td>
<td>0.64</td>
<td>0.60</td>
<td>0.55</td>
<td>0.48</td>
<td>0.43</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>LRSQ(3,3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.33</td>
<td>0.71</td>
<td>1.08</td>
<td>1.63</td>
<td>1.98</td>
<td>2.27</td>
</tr>
<tr>
<td>Vol</td>
<td>0.64</td>
<td>1.34</td>
<td>2.12</td>
<td>3.62</td>
<td>4.88</td>
<td>6.31</td>
</tr>
<tr>
<td>SR</td>
<td>0.51</td>
<td>0.53</td>
<td>0.51</td>
<td>0.45</td>
<td>0.41</td>
<td>0.36</td>
</tr>
</tbody>
</table>
Table 5: Conditional expected excess swap returns.
The table reports results from regressing nonoverlapping monthly excess swap returns on previous month’s term structure slope and implied volatility (including a constant). Consider the results for the 5-year maturity: The excess return is on an interest rate swap with a 1-month forward start and a 5-year swap maturity. The term structure slope is the difference between the 5-year swap rate and 3-month LIBOR. The implied volatility is the normal implied volatility for a swaption with a 3-month option expiry and a 5-year swap maturity. Excess returns are in percent, and the term structure slopes and implied volatilities are standardized. The top panel shows results in the data, where each time series consists of 200 monthly observations from February 1997 to August 2013. $t$-statistics, corrected for heteroscedasticity and serial correlation up to 4 lags using the method of Newey and West (1987), are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. The lower panels show results in simulated data, where each time series consists of 600,000 monthly observations (50,000 years).
### Table 6: Changing characteristics of “level” factor at the zero lower bound.

The table shows the factor loadings of the first principal component of the term structure, when the short rate is away from the zero lower bound (ZLB, Panel A) and close to the ZLB (Panel B). The factor loadings are constructed from the eigenvector corresponding to the largest eigenvalue of the covariance matrix of weekly changes in swap rates. The first row in each panel shows results in the data, where the non-ZLB sample period consists of 620 weekly observations from January 29, 1997 to December 16, 2008, and the ZLB sample period consists of 246 weekly observations from December 16, 2008 to August 28, 2013. The following rows in each panel show results in simulated data, where each time series consists of 2,600,000 weekly observations (50,000 years). In the simulated data, the non-ZLB (ZLB) sample consists of those observations where 3-month LIBOR is larger (less) than 50 basis points.

<table>
<thead>
<tr>
<th></th>
<th>1 yr</th>
<th>2 yrs</th>
<th>3 yrs</th>
<th>5 yrs</th>
<th>7 yrs</th>
<th>10 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: non-ZLB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>0.53</td>
<td>0.48</td>
<td>0.43</td>
<td>0.36</td>
<td>0.31</td>
<td>0.27</td>
</tr>
<tr>
<td>LRSQ(3,1)</td>
<td>0.40</td>
<td>0.42</td>
<td>0.43</td>
<td>0.42</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>LRSQ(3,2)</td>
<td>0.40</td>
<td>0.41</td>
<td>0.42</td>
<td>0.42</td>
<td>0.41</td>
<td>0.38</td>
</tr>
<tr>
<td>LRSQ(3,3)</td>
<td>0.38</td>
<td>0.41</td>
<td>0.42</td>
<td>0.43</td>
<td>0.42</td>
<td>0.39</td>
</tr>
<tr>
<td><strong>Panel B: ZLB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>0.10</td>
<td>0.25</td>
<td>0.38</td>
<td>0.51</td>
<td>0.52</td>
<td>0.50</td>
</tr>
<tr>
<td>LRSQ(3,1)</td>
<td>0.17</td>
<td>0.25</td>
<td>0.33</td>
<td>0.44</td>
<td>0.52</td>
<td>0.58</td>
</tr>
<tr>
<td>LRSQ(3,2)</td>
<td>0.16</td>
<td>0.27</td>
<td>0.36</td>
<td>0.47</td>
<td>0.52</td>
<td>0.53</td>
</tr>
<tr>
<td>LRSQ(3,3)</td>
<td>0.13</td>
<td>0.24</td>
<td>0.33</td>
<td>0.47</td>
<td>0.53</td>
<td>0.56</td>
</tr>
</tbody>
</table>
Table 7: Degree of unspanned stochastic volatility.

The table reports $R^2$'s from regressing weekly changes in normal implied swaption volatilities on the first three principal components (PCs) of weekly changes swap rates. Squared PCs and a constant are also included in the regressions. The first row shows results in the data, where each time series consists of 866 weekly observations from January 29, 1997 to August 28, 2013. The following rows show results in simulated data, where each time series consists of 2,600,000 weekly observations (50,000 years).

<table>
<thead>
<tr>
<th></th>
<th>1 yr</th>
<th>2 yrs</th>
<th>3 yrs</th>
<th>5 yrs</th>
<th>7 yrs</th>
<th>10 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.11</td>
<td>0.14</td>
<td>0.17</td>
<td>0.19</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>LRSQ(3,1)</td>
<td>0.17</td>
<td>0.32</td>
<td>0.41</td>
<td>0.51</td>
<td>0.56</td>
<td>0.61</td>
</tr>
<tr>
<td>LRSQ(3,2)</td>
<td>0.12</td>
<td>0.22</td>
<td>0.25</td>
<td>0.21</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>LRSQ(3,3)</td>
<td>0.17</td>
<td>0.21</td>
<td>0.26</td>
<td>0.26</td>
<td>0.21</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Table 8: Level-dependence in volatility of 1-year swap rate.
The table reports results from regressing weekly changes in the normal implied swaption volatility at the 1-year swap maturity on weekly changes in the 1-year swap rate (including a constant). The first column shows results using data where the 1-year swap rate is below 0.08. The second to sixth column shows results conditional on the 1-year swap rate being in the intervals 0-0.01, 0.01-0.02, 0.02-0.03, 0.03-0.04, and 0.04-0.08, respectively. The top panel shows results in the data, where each time series consists of 866 weekly observations from January 29, 1997 to August 28, 2013. t-statistics, corrected for heteroscedasticity and serial correlation up to 4 lags using the method of Newey and West (1987), are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. The lower panels show results in simulated data, where each time series consists of 2,600,000 weekly observations (50,000 years).

<table>
<thead>
<tr>
<th></th>
<th>&lt;0.08</th>
<th>0.00-0.01</th>
<th>0.01-0.02</th>
<th>0.02-0.03</th>
<th>0.03-0.04</th>
<th>0.04-0.08</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.18***</td>
<td>1.20***</td>
<td>0.54***</td>
<td>0.28***</td>
<td>−0.02</td>
<td>−0.04</td>
</tr>
<tr>
<td></td>
<td>(2.80)</td>
<td>(7.52)</td>
<td>(3.36)</td>
<td>(3.50)</td>
<td>(−0.20)</td>
<td>(−0.62)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.05</td>
<td>0.52</td>
<td>0.25</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>LRSQ(3,1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.14</td>
<td>0.66</td>
<td>0.23</td>
<td>0.07</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.12</td>
<td>0.52</td>
<td>0.38</td>
<td>0.10</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>LRSQ(3,2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.12</td>
<td>0.58</td>
<td>0.32</td>
<td>0.18</td>
<td>0.09</td>
<td>−0.02</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.10</td>
<td>0.42</td>
<td>0.36</td>
<td>0.27</td>
<td>0.12</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>LRSQ(3,3)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.15</td>
<td>0.56</td>
<td>0.34</td>
<td>0.20</td>
<td>0.12</td>
<td>0.06</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.16</td>
<td>0.41</td>
<td>0.39</td>
<td>0.31</td>
<td>0.17</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Figure 1: Data and fit
Panel A1 shows time series of the 1-year, 5-year, and 10-year swap rate. Panel B1 shows time series of the normal implied volatilities on 3-month options on the 1-year, 5-year, and 10-year swap rate. Panels A2 and B2 shows the fit to the swap rates and swaption implied volatilities in case of the $LRSQ(3,3)$ specification. Panels A3 and B3 shows time series of the root mean squared pricing errors (RMSE) of swap rates and swaption implied volatilities, respectively. The units in Panels B1, B2, A3, and B3 are basis points. The grey areas mark the two NBER-designated recessions from March 2001 to November 2001 and from December 2007 to June 2009, respectively. Each time series consists of 866 weekly observations from January 29, 1997 to August 28, 2013.
Figure 2: Estimated factors
The figure displays time series of the estimated factors. The first, second, and third column shows the factors of the $LRSQ(3,1)$, $LRSQ(3,2)$, and $LRSQ(3,3)$ specification, respectively. The first row displays $Z_{1,t}$ and $U_{1,t}$. The second row displays $Z_{2,t}$ and possibly $U_{2,t}$. The third row displays $Z_{3,t}$ and possibly $U_{3,t}$. The thin black lines show the term structure factors, $Z_{1,t}$, $Z_{2,t}$, and $Z_{3,t}$. The thick grey lines show the unspanned stochastic volatility factors, $U_{1,t}$, $U_{2,t}$, and $U_{3,t}$. The grey areas mark the two NBER-designated recessions from March 2001 to November 2001 and from December 2007 to June 2009, respectively. Each time series consists of 866 weekly observations from January 29, 1997 to August 28, 2013.
Figure 3: Level-dependence in volatility of the term structure factors
For each term structure factor, its instantaneous volatility is plotted against its level.
The first, second, and third column correspond to the $LRSQ(3,1)$, $LRSQ(3,2)$, and $LRSQ(3,3)$ specification, respectively. The first, second, and third row correspond to $Z_{1,t}$, $Z_{2,t}$, and $Z_{3,t}$, respectively. Each plot contains 866 weekly observations from January 29, 1997 to August 28, 2013. The grey areas mark the possible range of factor volatilities for a given factor level.
Figure 4: Conditional density of short rate
Conditional on the short rate being between 0 and 25 basis points, Panels A–C show the distribution of the short rate at a 1-year, 2-year, and 5-year horizon, respectively. Panel D shows the mean and modal path of the short rate. The distributions are obtained from 2,600,000 weekly observations (50,000 years) of the short rate simulated from the $LRSQ(3,1)$ specification.
Figure 5: Level-dependence in volatility of 1-year swap rate
The figure shows the normal implied volatility of the 3-month option on the 1-year swap rate (in basis points) plotted against the 1-year swap rate.
Figure 6: Level-dependence in volatility of 1-year swap rate.
The figure shows weekly changes in the normal implied volatility of the 3-month option on the 1-year swap rate (in basis points) against weekly changes in the 1-year swap rate, conditional on the 1-year swap rate being in the intervals 0-0.01 (Panel A), 0.01-0.02 (Panel B), 0.02-0.03 (Panel C), and 0.03-0.04 (Panel D). Black crosses denote the data and grey dots denote data simulated from the $LRSQ(3,2)$ specification.
References


65