Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities?*

Scott Joslin†

First Draft: February 12, 2006
Current Draft: November 14, 2006

*I am very grateful to Ken Singleton for numerous discussions and comments. I also thank Caio Almeida, Snehal Banerjee, Darrell Duffie, Damir Filipovic, Jeremy Graveline, Kris Jacobs, Yaniv Konchitchki, and Ilya Strebulaev for helpful discussions.

†Stanford Graduate School of Business, 518 Memorial Way, Stanford, CA 94305. Phone: 650.725.6858. joslin@stanford.edu
Abstract

In fixed income markets, volatility is unspanned if volatility risk cannot be perfectly hedged with bonds. I first show that a broad class of affine term structure models can be drift normalized and show when the standard variance normalization can be obtained. Using this normalization, I find general conditions for affine term structure models to exhibit unspanned stochastic volatility (USV). I show that the USV conditions restrict both the mean reversions of risk factors and the cross section of conditional yield volatilities. The restrictions imply that previously studied affine USV models are incapable of generating the observed cross section of yield volatilities. However, more general USV models can match the cross section of bond volatilities.

KEYWORDS: Unspanned stochastic volatility, affine term structure models
Introduction

Bond markets are incomplete if there are sources of risk that affect fixed income derivatives that cannot be hedged with bonds. Such incompleteness might arise, for example, if bond prices depend on the volatility of the short-term interest rate due to convexity effects and if expectations of future interest rates depend on volatility. These two effects can exactly cancel out in which case bond prices will be independent of volatility. Collin-Dufresne and Goldstein (2002) show that such cancellation is possible in terms structure models with three or more risk factors. Because volatility clearly affects bond option prices, fixed-income options can complete the market. They refer to this form of bond market incompleteness as unspanned stochastic volatility (USV).

A growing literature finds empirical support for separate risk factors driving fixed income derivative prices. Collin-Dufresne and Goldstein (2002) find evidence for the existence of unspanned volatility by showing in a regression analysis that returns for cap straddles, which are particularly sensitive to volatility, are only slightly correlated with changes in bond prices. Heidari and Wu (2003) find that although level, slope, and curvature factors explain nearly all of the variation in yields, they explain only a fraction of the variation in swaption implied volatilities. Using away-from-the-money interest rate cap price data, Li and Zhao (2006) find that a quadratic affine term structure model estimated on interest rate swaps cannot hedge cap straddles well. Andersen and Benzoni (2005) find factors driving short-horizon volatility that are unrelated to yield curve movements by analyzing high frequency data.

In this paper, I examine the ability of affine term structure models with incomplete bond markets to explain the dynamics of the cross section of interest rates.\(^1\) The literature has come to conflicting conclusions regarding the
ability of low-dimensional dynamic term structure models to jointly capture both the USV phenomenon and the conditional term structure of volatility. Thompson (2004) empirically investigates the ability of affine models with a single volatility factor to match the dynamics of yields. The estimated models he considers, with and without USV imposed, match the dynamics of longer maturities but the conditional volatility of the short rate is not matched. In contrast, Collin-Dufresne et al. (2006a) analyze similar models and find that their estimated three and four factor USV models capture the dynamics of the short rate while non-USV models do not. However, they do not fully examine the dynamics of longer maturities. Bikbov and Chernov (2005), using futures and short dated options on short maturities, find that their USV model substantially misprices the options.

In this paper, I show that the apparent failure of these USV models to simultaneously match the volatilities of long- and short-maturity bonds, and the implied volatilities of bond options, can be explained by the restrictions on the relative rates of mean reversion of risk factor implied by USV. To illustrate these issues taken up more formally below, consider the swap-implied yields for the period April 1987 to September 2005 displayed in Figure 1 (see Section 3 for a full description of the data). Evident is the widely studied level factor in yields. The fact that a level factor explains much of the variation in yields implies the existence of a persistent factor with very slow mean reversion driving the short rate. Figure 2 shows the time series of yield conditional volatilities and demonstrates how the conditional volatilities of longer maturity yields move together. This implies that the persistent factor driving the short rate has stochastic volatility. On the other hand, short maturity conditional volatilities have greater variability than longer maturity yields, suggesting the existence of a quickly mean reverting factor driving the short rate that has stochastic volatility.

Any factor with stochastic volatility will generate a convexity effect. For
USV to be present, this effect must cancel across all maturities so that this volatility does not affect bond prices. I show that, since the convexity effect involves a quadratic term, the cancellation of a risk factor’s convexity effect can only be due to volatility affecting expectations of another factor with twice the level of mean reversion. With a single volatility factor and two or three non-volatility factors, as studied in the prior literature, it is impossible to cancel the convexity effects of both a quickly mean reverting factor and a slowly mean reverting factor. This creates the tension seen in the prior empirical results. If USV is maintained, there can either be a level factor with stochastic volatility or a short end factor with stochastic volatility, but not both. Therefore, it is theoretically impossible to simultaneously match the volatilities of yields at both ends of the yield curve and option volatilities in these models.

In light of this tension, I then investigate if more generally USV affine models are capable of matching the cross section of yields and yield conditional volatilities. Before examining incompleteness conditions in affine models, I first construct a drift normalized canonical form. Cheridito et al. (2006) show that there exists affine term structure models which do not fall into the class studied by Dai and Singleton (2000). I show that all affine models with state space $\mathbb{R}^+_M \times \mathbb{R}^{N-M}$ can be uniquely drift normalized. Additionally, I characterize when it is not possible to use the alternative variance normalization of Dai and Singleton (2000).

With a more general characterization of affine term structure models in hand, I then examine conditions to exhibit incomplete markets. This extends and generalizes the results of Collin-Dufresne and Goldstein (2002). I show that any incompleteness must involve stochastic volatility factors. This means that Gaussian affine term structure models, which have constant volatility, cannot generate incomplete markets. I classify all models with incomplete markets when there is a single stochastic volatility factor driving interest
rates and derive several USV specifications not addressed in the prior literature. In order to capture the dynamics of the cross section of yields, I analyze affine incomplete markets models with two factors driving volatility. Adding the additional volatility factor allows the incomplete markets models to potentially explain the cross section of yields and yield volatilities.

The rest of the paper is organized as follows. Section 1 presents the class of identified affine short rate models that is used for examining imposing incomplete markets and derives several general results regarding incompleteness in affine models. Section 2 provides an interpretation of the constraints derived in Section 1. Section 3 examines in more depth the ability of existing affine USV models to explain both the cross section of yields and yield volatilities. Section 4 derives conditions for incomplete markets in affine models with multiple factors driving volatilities. Finally, Section 5 concludes.

1 Incomplete Markets in Affine Models

In an affine short rate model, the short rate, \( r_t \), is driven by an \( N \)-dimensional Markov state variable, \( X_t \), such that

\[
\begin{align*}
  r_t &= \rho_0 + \rho_1 \cdot X_t, \\
  dX_t &= \mu_t dt + \sigma_t dB_t^Q, \\
  \mu_t &= K_0 + K_1 X_t, \\
  \sigma_t \sigma_t' &= H_0 + H_1 \cdot X_t,
\end{align*}
\]

where \( H_1 \cdot X_t = \sum_{i=1}^N H_i^t X_t^i \), an \( N \times N \) matrix.

Any claim with payoff at time \( T \) given by \( f(X_T) \) can be priced by the dis-
counted risk-neutral expected value

\[ E_t^Q \left[ e^{-\int_0^T r_s \, ds} f(X_T) \right]. \]

In this setting, Duffie and Kan (1996) show that zero coupon bond prices are given by

\[ P^T (X_t, t) = e^{A(T-t)+B(T-t) \cdot X_t}, \]

where the loadings \( A \) and \( B \) satisfy the Riccati differential equations

\[
\dot{B} = -\rho_1 + K_1^T B + \frac{1}{2} B^T H_1 B, \quad B (0) = 0,
\]

\[
\dot{A} = -\rho_0 + K_0^T B + \frac{1}{2} B^T H_0 B, \quad A (0) = 0.
\]

In the event that the set \( \{B(\tau)\}_{\tau \in \mathbb{R}^+} \) is contained in a subspace of \( \mathbb{R}^N \) of dimension less than \( N \), there will be incomplete bond markets. That is, there will be risks which affect fixed income derivatives but cannot be hedged using only bonds. Collin-Dufresne and Goldstein (2002), hereafter CG, show that it is possible in an \( A_1(3) \) model for volatility to not directly affect bond prices. Specifically, there is a rotation of the state variables so that the volatility of the short rate is one of the factors and its loading is zero.

Dai and Singleton (2000), hereafter DS, point out two key issues that arise in applying affine term structure models. The first is econometric identification. Since the state variable is unobserved, it is possible that two distinct models may be observationally equivalent. A second issue is admissibility. In order to have a well-defined process we must guarantee, for example, that \( \sigma_t \sigma_t' \) is positive semi-definite for all possible states. It will be convenient to first impose admissibility and identification conditions before imposing incompleteness conditions. Since any affine USV model is a restricted version of a general affine model, this is a natural order for applying the types of
In order to guarantee an admissible and identified term structure model, I proceed differently from DS. First, I consider the class of admissible regular affine diffusion with state space $\mathbb{R}_+^M \times \mathbb{R}^{N-M}$ given in Duffie et al. (2003). Cheridito et al. (2006) show that this class is strictly larger than class considered by DS when $M \geq 2, (N - M) \geq 2$. Second, in contrast to DS who identify the models by normalizing the variance structure, I normalize the drift structure.

Define an affine $Q$-model to be drift normalized when it can be written in the form:

$$
K_1 = \begin{bmatrix} K_V & 0 \\ K_V G & K_G \end{bmatrix}, \quad H_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_0^G \end{bmatrix}, \quad H_1^i = \begin{bmatrix} \Sigma_i^V & 0 \\ 0 & \Sigma_i^G \end{bmatrix}, \quad i \leq M
$$

where the matrix $K_G$ is diagonal with entries strictly increasing on the diagonal. The $\Sigma_i^V$ are $M \times M$ matrices which are all zero except the $(i, i)$-entry, which is 1, and $\Sigma_i^G$ are positive semi-definite matrices. Additionally, $K_{0,i} \geq 0$ for $i \leq M$, $H_1^i$ is a matrix of zeros for $i > M$ and $\rho_{1,i} = 1$ for $i > M$. 4

I will refer to the first $M$ factors as volatility factors and the last $N - M$ factors as conditionally Gaussian factors, since conditional on the path of the first $M$ factors they will be normally distributed. For complete details of the identification see Appendix A.

The uniqueness of this representation is guaranteed by

**Theorem 1.1.** Every non-degenerate affine model with state space $\mathbb{R}_+^M \times \mathbb{R}^{N-M}$ is equivalent to exactly one model in drift normalized form.

The models of DS take the form of (1) with $\Sigma_i^G$ diagonal for $i = 0, 1, \ldots, M$, but $K_G$ is unconstrained. It is always possible to simultaneously diagonalize two or fewer matrices, but in general it is not possible to simultaneously
diagonalize three or more matrices. Theorem 1.2 characterizes more precisely the relationship between drift normalized models and the variance normalized models of DS.

**Theorem 1.2.** A drift normalized model is equivalent to a DS normalized model if and only if $\Sigma_0$ is strictly positive definite and there is a single set of vectors $\{x_1, \ldots, x_{N-M}\} \subset \mathbb{R}^{N-M}$ which are all eigenvectors of $\Sigma_0^{-1}\Sigma_1^G, \ldots, \Sigma_0^{-1}\Sigma_M^G$.

Cheridito et al. (2006) prove that any affine term structure model with state space $\mathbb{R}_+^M \times \mathbb{R}_{N-M}$ may be DS-normalized when $M \leq 1$ or $M \geq N-1$. Since the hypothesis of Theorem 1.2 trivially holds in these cases, this generalizes their result. Additionally, the converse of Theorem 1.2 directly proves that the $A_2(4)$ example they consider cannot have its variance diagonalized since in their example $\Sigma_0^{-1}\Sigma_1^G$ and $\Sigma_0^{-1}\Sigma_2^G$ have different distinct eigenvectors.

My approach of imposing admissability prior to an examination of USV differs from the approach in CG and Collin-Dufresne et al. (2006b). There admissability is imposed after constraining the model to achieve incompleteness. An advantage of imposing admissability first is that many higher factor incomplete market models are immediate from lower factor models. As long as the new factors do not affect the existing loadings, either by entering through the conditional mean or convexity terms, markets will still be incomplete. For instance, if the parameters in (1) are for an $A_1(N)$ model with incomplete markets, then an $A_1(N+1)$ incomplete markets model can be obtained by:

$$K_1 \Rightarrow \tilde{K}_1 = \begin{bmatrix} K_V & 0 & 0 \\ K_{VG} & K_G & K^*_G,1 \\ 0 & 0 & K^*_G,2 \end{bmatrix}, \quad H_1^1 \Rightarrow \tilde{H}_1^1 = \begin{bmatrix} H_1^1 & 0 \\ 0 & 0 \end{bmatrix}.$$  

Since $H_0$ doesn’t enter into the Riccati equation for the factor loadings, $B$,
$H_0$ is not restricted except by admissability. Although derived differently, the $A_1(3)$ and $A_1(4)$ models in Collin-Dufresne et al. (2006a) are related in this manner, and so are special cases of my construction.

Normalizing so that $K_G$ is diagonal has several practical side benefits for empirical estimation. First, this parametrizes the model in terms of the eigenvalues which are invariant to rotation. This allows for a direct comparison of models estimated on different datasets or even estimated with a different number of factors. Second, there is a natural prior for the eigenvalues. Even without estimation, one can be reasonably confident that one eigenvalues is near zero with others corresponding to shorter half-lives from about six months to ten years. Further, constraints on the eigenvalues, such as $Q$–stationarity, become linear and therefore easy to impose.

Presently, there are two advantages to diagonalizing the drift. First, the Riccati equations for the last $N - M$ factors are uncoupled from the first $M$ equations. Second, by diagonalizing $K_G$, the last $N - M$ ricatti equations are uncoupled and have simple closed form solutions. The last $N - M$ Riccati equations are

$$\dot{B}_j = -\rho_{1,j} + K_{jj} B_j, \quad B_j(0) = 0.$$  

The solution is

$$B_j(\tau) = \frac{\rho_j}{K_{1,jj}} (1 - e^{K_{1,jj} \tau}).$$

In the non-diagonalizable case, there are also closed from solutions which are exponential polynomials. It is easy to see conditions under which exponential polynomials are colinear:

**Lemma 1.3.** For $(c_i, d_i)$ distinct, if

$$\sum_{i=1}^{I} \alpha_i t^i e^{c_i t} = 0 \quad \forall t \geq 0$$

then $\alpha_i = 0$ for all $i$. 

10
Before applying this to the bond loadings, we must first rule out cases where there are extraneous factors which affect neither bond prices nor fixed income derivative prices. For example, suppose that $X_t$ is a 2-dimensional Brownian motion with independent components and $r_t = X^1_t + X^2_t$. In this case, $r_t$ is itself a 1-dimensional Brownian motion. The factor $X^1_t - X^2_t$ represents a source of uncertainty which is independent of the short rate and which does not affect any prices.

**Theorem 1.4.** Necessary and sufficient conditions that an $A_M(N)$ model in the form of (1) does not reduce to an $A_M(N-1)$ model are that there are no eigenvectors of $K_G$ orthogonal to $\rho_G$. In particular, the geometric multiplicity of each eigenvalue of $K_G$ must be 1 and each eigenvalue corresponds to only one Jordan block.

The proof of Theorem 1.4 gives the important result:

**Theorem 1.5.** In any non-degenerate $A_M(N)$ in the form of (1), $\{X_{M+1}, \ldots, X_N\}$ are always spanned. That is, for any $\{c_n\}$ it cannot be true that $\sum_{n=M+1}^N c_n B_n(\tau) = 0, \forall \tau$ unless all $c_n = 0$.

from which it follows:

**Corollary 1.6.** Bond markets are complete in any $A_0(N)$ model.

CG prove this corollary for the special cases $N = 2$ and $N = 3$ by explicitly examining a Taylor series expansion of the factor loadings.

Theorem 1.5 shows that that any incompleteness must involve the stochastic volatility factors. Since the conditionally Gaussian factors’ loadings are known in closed form, this greatly simplifies the problem of identifying affine models with incomplete markets.

It is particularly easy to characterize incomplete markets in $A_1(N)$ models.
Lemma 1.7. For any $N \geq 3$, there exist $A_1(N)$ models with incomplete markets. When in the form of (1), the loading on the volatility factor will be an exponential polynomial of the form

$$B_1(\tau) = \sum_i \alpha_i \tau^{n_i} \text{Real}(e^{k_i \tau})$$

where $k_i$ may be complex if there are complex eigenvalues of $K_G$.

For fixed $N$, there will be several distinct classes of USV models. A general procedure to find all $A_1(N)$ models with incomplete markets is as follows. First, consider the possible Jordan decompositions for $K_G$. Enumerating the possible Jordan block decompositions guarantees a complete classification. For a fixed Jordan block structure of $K_G$, we can then compute closed form solutions for $B_G$. The Riccati equation for $B_1$ is:

$$\dot{B}_1 = \sum_{j=1}^{N} K_{1,j} B_j + \frac{1}{2} B_1^2 + B_G^\top \Sigma_1^G B_G - \rho_{1,1} \quad (2)$$

Since $B_1$ is a linear combination of the other loadings, each of these terms will be an exponential polynomial. We can then apply Lemma 1.3 and equate coefficients of each monomial term.

To illustrate this procedure, consider the case of an $A_1(4)$ model. For a general drift-normalized $A_1(4)$ model, the conditional means and variances take the form:

$$\text{drift}(X_t) = K_0 + K_1 X_t, \quad \text{var}(X_t) = H_0 + H_1^1 X_t^1.$$

The diagonal entries of $K_1$ control the rates of mean reversion of the risk factors. For example, if $K_{1,44} = \log(\frac{1}{2})/10 \approx -0.069$, there will be a relatively persistent risk factor with a half-life of 10 years. This persistent risk factor
will have stochastic volatility provided \( H_{1,44} > 0 \).

\[
K_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
* & * & 0 & 0 \\
* & 0 & * & 0 \\
* & 0 & 0 & \approx 0 \\
\end{bmatrix}, \quad H_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
0 & * & * & * \\
0 & * & * & * \\
0 & * & * & + \\
\end{bmatrix}
\]

In this case, the persistent risk factor will generate a quadratic convexity effect decaying at rate \(-2K_{1,44}\). For this convexity effect to possibly cancel, an eigenvalue will need to be restricted to be \(2K_{1,44}\).

\[
K_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
* & * & 0 & 0 \\
* & 0 & 2K_{1,44} & 0 \\
* & 0 & 0 & K_{1,44} \\
\end{bmatrix}, \quad H_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
0 & * & * & * \\
0 & * & * & * \\
0 & * & * & + \\
\end{bmatrix}
\]

Restricting \(K_{1,33} = 2K_{1,44}\) allows the possibility that the convexity effect of the persistent factor generated by \(H_{1,44}\), but other positive entries in \(H_1\) will generate convexity effects as well. There are two possibilities. One case is that \(K_{1,22}\) could be unrestricted and then only the persistent Gaussian factor generating a convexity effect. This requires zeroing all but one entry in the lower 3 × 3 matrix of \(H_1\) to be zero:

\[
K_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
* & * & 0 & 0 \\
* & 0 & 2K_{1,44} & 0 \\
* & 0 & 0 & K_{1,44} \\
\end{bmatrix}, \quad H_1 = \begin{bmatrix}
* & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & + \\
\end{bmatrix}
\]

This is the case given in Collin-Dufresne et al. (2006b). Further restriction on \(K_{1,11}, K_{1,31}, K_{1,41}\) are needed to ensure that volatility affects the conditional expectations in such as way as to cancel the convexity effects. The exact
form of the restriction can be derived substituting the closed form loading into the Riccati equation and applying Lemma 1.3. Appendix C contains details of this derivation for the alternative case where $K_{1,22}$ is restricted to be twice $K_{1,33}$ which allows the third factor to have stochastic volatility and generate a convexity effect.

It is apparent that in either case not only are the mean reversions of the risk factors restricted, but the covariance structure is required to be very simple. Section 3 illustrates that the simple covariance structure results in a simple cross section of conditional yield volatilities.

### 2 Interpreting USV Constraints

There are three general types of constraints. The first type of constraints are restrictions on the rates of mean reversion of the risk factors. Second, the conditional covariances of the risk factors are constrained in order to not introduce convexity effects which cannot be cancelled. These two conditions allow for the possibility of expectations effects to exactly cancel any convexity effects generated across maturities. The third type of constraints requires volatility to affect expectations in such a way as to provide the cancellation.

For illustration, consider the case of an $A_1(N)$ model with distinct eigenvalues in the drift matrix and the identification and admissibility constraints of Section 1 imposed. The sensitivity of the log price of a $\tau$-year to maturity bond to a shock in a Gaussian risk factor is proportional to $(1 - e^{-\kappa \tau})$, where $\kappa$ is the rate of mean reversion of the factor. As the maturity grows, bond prices become less sensitive to shocks in the Gaussian risk factors. The sensitivity of bond prices to the volatility factor is more complicated, representing a mix of expectations and convexity effects. Because the convexity effect is quadratic, the convexity effect generated due to stochastic volatility of a factor mean
reverting at rate $\kappa$ can only be cancelled by a factor mean reverting at rate $2\kappa$. More generally, for two factors which mean revert at rates $\kappa_1$ and $\kappa_2$, the convexity effect generated by stochastic covariance between the factors can only be cancelled by volatility affecting expectations of a factor which mean reverts at rate $\kappa_1 + \kappa_2$.

Convexity effects also introduce restrictions on the covariance structure of the risk factors. The local covariance of the conditionally Gaussian risk factors are affine in the volatility factor, $\text{var}(X_G) = \Sigma_0^G + \Sigma_1^G X_V$, and each non-zero entry of $\Sigma_1^G$ generates a convexity effect. Since each convexity effect places a restriction on the rate of mean reversion of one of the risk factors, some factors must have constant volatility. For example, the risk factor which mean reverts most quickly must have constant volatility because its quadratic convexity effect will decay twice as fast any expectation effect. Each non-zero element of $\Sigma_1^G$ will require there to be a risk factor to mean revert at a specific rate to cancel the convex effect. Therefore, imposing $n$ restrictions on the rates of mean reversion of the risk factors will allow $n$ entries in the $(N - 1) \times (N - 1)$ matrix $\Sigma_1^G$ to be non-zero. That is, only by adding restriction to the factor mean reversion rates can the convexity effects of stochastic covariance be cancelled.

Finally, constraints that ensure cancellation of the expectations and convexity terms to cancel must be imposed. This is only possible provided the first two types of restrictions hold. These restriction are represented in restrictions on the effect of the volatility factor on the conditional means of the Gaussian state variables through the parameters $K_{11}$. 

15
3 The Cross Section of Bond Yields

A robust finding of Litterman and Scheinkman (1991) is that changes in yields are explained by three principal components: a level, slope, and curvature factor. In an affine term structure model, the PC loadings will be, at least approximately, a linear combination of the yield loadings \( B_y(\tau) \equiv -B(\tau)/\tau \). To see this, consider a fixed collection of maturities \( \vec{\tau} = (\tau_1, \tau_2, \ldots, \tau_n) \). Suppose the factors are rotated and scaled so that \( \| B_y(\vec{\tau}) \|_2 = 1 \) and the covariance matrix of \( X_t, \Sigma \), is diagonal with eigenvalues decreasing on the diagonal. The covariance matrix of the yields is then

\[
B_y(\vec{\tau})^\top \Sigma B_y(\vec{\tau}).
\]

This is exactly the decomposition which PC analysis seeks to achieve and so \( B_y(\vec{\tau}) \) must give the model implied PC loadings.

In order to have a level factor, a linear combination of the loading must be relatively constant across maturities. For an \( A_1(N) \) USV model with distinct real eigenvalues, this requires an eigenvalue to be approximately zero. That is, one factor must have very slow mean reversion. Dai and Singleton (2000), Duffee (2002), Duarte (2003), and other authors have found such a near unit root under a variety of affine specifications. For an \( A_1(3) \) USV model, since one eigenvalue is constrained and one factor does not affect yields, there is essentially only a single free eigenvalue. This means having a level PC essentially constrains all the loadings. An \( A_1(N) \) USV model has at most \( N - 2 \) free eigenvalues of \( K^G_1 \). The more flexible the volatility structure (i.e., the more convexity effects that are generated), the fewer free eigenvalues there will be.

Consider now the term structure of conditional volatility. In an affine model,
the local conditional variance of a $\tau$-year bond yield is

$$\text{var}_t(y_\tau) = \text{var}_t(B_y(\tau) \cdot X)$$

$$= B_y(\tau)\text{T} \text{var}_t(X)B_y(\tau)$$

$$= B_y(\tau)\text{T} H_0 B_y(\tau) + \sum_{i=1}^{M} B_y(\tau)\text{T} H_i B_y(\tau) X_i$$

The restrictions on the conditional means of the factors will enter the volatility through the $B_y$ loadings. Incomplete markets give no restrictions on $H_0$. This means that generally USV models will be capable of capturing the unconditional term structure of yield volatilities. An exception is the $A_1(3)$ USV model. If the only free eigenvalue of $K^D_1$ is approximately zero, agreeing with the first principal component of yield changes, the unconditional term structure of volatility will be approximately flat. In an $A_1(N)$ model the changes in conditional variances will be linear in the volatility factor:

$$\Delta \text{var}_t(y_\tau) = B_y(\tau)\text{T} H_1 B_y(\tau) \Delta X_1$$

Both the loadings and $H_1$ will be restricted by USV.

To examine the restrictions on the term structure of volatility empirically, I construct a time series of changes in volatility using 3 month LIBOR, 6 month LIBOR, and zero coupon yields up to 10 years bootstrapped from swaps rates. Weekly data was obtained from Datastream for the period April 1987 to October 2005. From the data, univariate EGARCH(1,1) models were estimated for each time series to obtain estimates of the changes in conditional variances. Time series for the EGARCH volatilities are shown in Figure 2. The most striking feature is that the volatilities of maturities beyond 2 years move virtually together. The loadings for the first principal component, which explained 67.5% of the variation, are shown in Figure 4. The principal component loadings decline very quickly until 2 years and once
past 5 years are practically flat.

The fraction of variance explained by the first principal component heuristically represents an upper bound on the ability of an $A_1(N)$ model to explain variation in the conditional variances of yields, since such model have only a single factor driving volatility. An unconstrained $A_1(3)$ model allow for generally flexible variation in the cross section of conditional volatility. Consider the case where $\rho_{1,1} = 0$, making volatility locally uncorrelated with the short rate, and constant correlation between the second and third risk factors $H^1_{1,23} = 0$, then

$$
\Delta \text{var}_t(y_\tau) \approx \left( B^2_{y_1}(\tau)H^1_{1,22} + B^2_{y_3}(\tau)H^1_{1,33} \right) \Delta X_1
$$

Figure 5 plots this for various values of the parameters. Note that by having a risk factor which is very persistent and a risk factor which is more transient, both with stochastic volatility, the model implied first principal component can be quite close to the observed first principal component.

In contrast, $A_1(3)$ and $A_1(4)$ USV models necessarily have much simpler variation in the cross section of conditional variances. For the $A_1(3)$ and $A_1(4)$ USV models in CGJ, $H^1_1$ will be of the form

$$
H^1_1 = \begin{bmatrix}
1 & 0 & 0 \\
0 & \beta_{12} & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad H^1_1 = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \beta_{12} & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.
$$

Because many of the entries in $H^1_1$ are constrained to be zero to not introduce convexity effects, only the first two loadings enter into changes in conditional volatilities. The fact that also $B_1 \propto B_2$ implies that in these cases

$$
\Delta \text{var}_t(y_\tau) \propto \frac{(1 - e^{a\tau})^2}{(a\tau)^2} \Delta X_t.
$$
This is inconsistent with the cross section of conditional volatilities for any $a$. For example, the estimates in Table 2A of Collin-Dufresne et al. (2006a) imply a half-life of 8.0 years for the $A_1(4)$ USV model. This value is consistent with a level factor in yield changes. Comparing to Figure 3, we see that the model implies that counterfactually volatilities change according to a level factor also. However, if the factor with stochastic volatility had faster mean reversion, longer maturity yields would have nearly constant volatility.

A principal components analysis of maturities one year and beyond using either EGARCH volatilities or short dated swaption implied volatilities shows that indeed a level factor explains most of the variation of changes in variances. If one is only interested in modelling these maturities, an $A_1(4)$ model with incomplete markets could be satisfactory.\footnote{Incomplete Market Models with Multiple Volatility Factors}

\section{Incomplete Market Models with Multiple Volatility Factors}

In order to capture the term structure of volatility within an affine USV setting, we can consider an $A_2(4)$ model. The methods developed in Section 1 can be used to find conditions for incomplete markets in $A_2(N)$ models. As noted previously, every equivalence class of $A_1(N-1)$ USV model immediately generates an $A_2(N)$ USV model by the addition of an additional volatility factor which doesn’t affect the existing factor loadings.

\[
K_1 = \begin{bmatrix}
0^* & 0 & 0 & 0 \\
0 & K_{21}^* & 0 & 0 \\
0 & 0 & K_{31}^* & K_{33}^* \\
0 & 0 & 2K_{33}^* & K_{41}^* & K_{42}^* & 0 \\
\end{bmatrix}, \quad \Sigma_1^G \text{ free}, \quad \Sigma_2^G = \begin{bmatrix} \beta_{23} & 0 \\
0 & 0 \end{bmatrix}
\]
Where the asterisks denote the restricted parameters, with the same restrictions as an $A_1(3)$ USV model. $X_1$ will represent a spanned volatility factor while $X_2$ will represent an unspanned volatility factor.

The restriction on the eigenvalues and which factors are allowed stochastic volatility will still be important in matching the cross section of yields and yield volatilities. When $K_{33} \approx 0$, there will be a level factor in both yields and yield volatilities. So if $X_1$ represents a short end factor, this will allow for the observed difference in short end volatility. In this case, the level volatility factor will be unspanned while the short end volatility factor will be spanned.

Another class of $A_2(N)$ incomplete market models have volatility loadings of the form $c_0 + c_1 e^{k\tau/2}$, where $k$ is an eigenvalue of $K_G$. This allows the quadratic terms in the volatility factors’ differential equations to cancel. Incomplete markets requires $\alpha \cdot B(0) = 0$ and $\alpha \cdot \dot{B}(0) = 0$ for some $\alpha$. This implies that either the loadings for $B_1$ satisfies a quadratic equation in $B_3, \ldots, B_N$ (which are known in closed form due to the Jordan block rotation) or several terms must cancel. This makes it unlikely that any cases other than the obvious cases mentioned above will work. For fixed $N$, the various possibilities can be considered. For $N = 3$, it can be verified by considering cases that only the obvious guess will work:

**Lemma 4.1.** There is a single equivalence class of DS-normalized $A_2(3)$ models which exhibit incomplete markets.

The bond loadings will be given by:

\[
B_i(\tau) = \frac{2\rho_i}{K_{3,3}} (1 - e^{K_{3,3} \tau/2}), \quad i = 1, 2
\]

\[
B_3(\tau) = \frac{\rho_3}{K_{3,3}} (1 - e^{K_{3,3} \tau})
\]
And the following constraints will hold:

\[
\begin{align*}
\beta_{1,3} &= 0 \\
\beta_{2,3} &= 0 \\
K_{3,1} &= \frac{2\rho_1^2}{\rho_3 K_{3,3}} \\
K_{1,1} &= -\frac{-\rho_1 K_{3,3}^2 + 2K_{2,1}\rho_2 K_{3,3} + 4\rho_1^2}{2\rho_1 K_{3,3}} \\
K_{2,2} &= -\frac{-\rho_2 K_{3,3}^2 + 2K_{1,2}\rho_1 K_{3,3} + 4\rho_2^2}{2\rho_2 K_{3,3}} \\
K_{3,2} &= \frac{2\rho_2^2}{\rho_3 K_{3,3}} \\
\rho_2 &\neq \rho_3 \quad (This \ ensures \ the \ model \ does \ not \ reduce \ to \ an \ A_1(2))
\end{align*}
\]

The additional constraints \( \beta_{1,1} = \beta_{2,2} = 1, \rho_3 = 1, K_{0,3} = 0 \), give an admissible identified \( A_2(3) \) USV model.

Another possibility would be to consider an \( A_1(5) \) USV model. This would allow for both a quickly mean reverting and a slowly mean reverting Gaussian risk factor with stochastic volatility. The other two Gaussian factors could provide cancellation of the convexity terms. The cross section of conditional variance will vary as in (3), and thus \( A_1(5) \) USV models have the ability to produce variation in the cross section of conditional variance as in Figure 5.

5 Conclusion

This paper investigate general conditions for incomplete markets in affine term structure models. I first derive conditions under which an regular affine term structure model with state space \( \mathbb{R}_+^M \times \mathbb{R}^{N-M} \) can be representing in
the variance normalized form of Dai and Singleton (2000). I show that any regular affine term structure model with state space $\mathbb{R}_+^M \times \mathbb{R}^{N-M}$ may be uniquely normalized by identifying the drift. While generally useful, this canonical form is particularly helpful for imposing incomplete market constraints due to the fact that it provides simple closed form solutions for some bond loadings.

I show that incomplete markets imposes strong restrictions on both the conditional mean and volatility of the factors in an affine term structure model. Typically the risk neutral mean reversion of one factor will be twice another in order for a quadratic convexity term to cancel. Since incompleteness by definition requires one factor to not affect bond prices, this restriction can be severe in low factor models. Incompleteness also restricts the number of factors which can have stochastic volatility. Coupled with the mean reversion restrictions, this creates a further tension in the models ability to match the cross section of conditional yields and yield volatilities. Including multiple volatility factors allows for affine USV models to potentially capture both the cross section of yield volatilities.
A General Affine Canonical Specification

This appendix presents a canonical model specification for general affine term structure models. The specification relies on rotating the drift structure rather than the variance structure as in DS. Here the drift of the Gaussian factors is typically given by a diagonal matrix and thus parametrized by the eigenvalues, a model invariant. This specification allows for the most general affine model with state space $\mathbb{R}_M^+ \times \mathbb{R}^{N-M}$ which is larger than the class of $A_M(N)$ models in DS when $M \geq 2, N-M \geq 2$. Additionally, by ordering the eigenvalues, a rotational indeterminacy present in DS is removed. Finally, as discussed in Section 1, the bond loadings for the Gaussian factors are simple exponential functions.

For the affine term structure model

\[
\begin{align*}
  r_t &= \rho_0 + \rho_1 \cdot X_t, \\
  dX_t &= \mu_t dt + \sigma_t dB_t, \\
  \mu_t &= K_0 + K_1 X_t, \\
  \sigma_t \sigma_t' &= H_0 + H_1 \cdot X_t,
\end{align*}
\]

where

\[
K_1 = \begin{bmatrix} K_V & 0 \\ K_{VG} & K_G \end{bmatrix}, \quad H_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma^G \end{bmatrix}, \quad H_i^i = \begin{bmatrix} \Sigma^V_i & 0 \\ 0 & \Sigma^G_i \end{bmatrix},
\]

with $K_V$ an $M \times M$ matrix, $K_G$ and $\Sigma^G_i (N-M) \times (N-M)$ matrices, and $\Sigma^V_i M \times M$ matrices, I define the drift normalized canonical representation as follows. For the parameters $\Theta = (\rho_0, \rho_1, K_0^P, K_1^P, K_0^Q, K_1^Q, C^G_{M+1}, C^G_{M+2}, \ldots, C^G_N)$, impose the constraints:

1. $K_G^Q$ is diagonal with entries increasing on the diagonal.\(^6\)
2. $C_i^G$ is lower triangular and gives the Cholesky factorization of $\Sigma_i^G$: $\Sigma_i^G = (C_i^G)^T (C_i^G)$.

3. $K_{0,n}^Q = 0$, $n > M$.

4. $\Sigma_{i,jk} = 1$, if $j = k = i$, or 0 otherwise.

5. $\rho_{1,n} = 1$, $n > M$.

6. $\rho_{1,n} < \rho_{1,n+1}$, $n < M$.

7. $K_{V,ij}^P > 0$, $K_{V,ij}^Q > 0$ if $i \neq j$.

8. $K_{0,n}^P \geq \frac{1}{2}$, $K_{0,n}^Q \geq \frac{1}{2}$, $n \leq M$.

Condition 3 normalizes the level of the Gaussian factors. Conditions 4 and 5 normalize the scales of the factors. Separate cases where $\Sigma_1^V = 0$ can be considered, but Lemma 1.4 shows that there are not cases where $\rho_{Gi} = 0$. Condition 6 removes a permutation indeterminacy in the volatility factors. Conditions 7 ensures admissability. Condition 8 enforces boundary non-attainment under both measures which relies on using the market price of risk in Cheridito et al. (2004). However, this condition could be modified if an alternative market price of risk is used.

Ordering the diagonal of $K_G$ removes a permutation indeterminacy. Conditions to remove a permutation-type indeterminacy are also needed but omitted in the normalization of DS as well. For example, any $A_0(N)$ model with their normalizations is equivalent to another normalized $A_0(N)$ model with the eigenvalues on the diagonal arbitrarily reordered. This is because generically the real Schur decomposition, of which lower triangular is a special case, is unique only up to reordering the diagonal blocks. Such a rotation can be achieved with a matrix $U$ formed by applying the Gram-Schmidt process to a set of eigenvectors in the desired ordering. The additional constraint of
\[ \kappa_{i,i} \leq \kappa_{i+1,i+1}, i > M, \] removes this permutation-type indeterminacy and normalizes their model. Similar ordering conditions on their \( \beta \) parameters are needed in the \( A_M(N) \) models with \( M \geq 1 \). Also, by considering real Schur decomposition with \( 2 \times 2 \) blocks corresponding to complex eigenvalues, their normalization can be easily expanded to consider the complex eigenvalue case.

Neglecting the special cases mentioned previously, every affine model with state space \( \mathbb{R}_+^M \times \mathbb{R}^{N-M} \) is observationally equivalent to exactly one such model. More formally, define a DFS affine term structure \( Q \)-model to be a parameter vector \( \Theta_X = (\rho_0, \rho_1, K_0, K_1, H_0, H_1) \) where \( r_t = \rho_0 + \rho_1 \cdot X_t \) and

\[
\begin{align*}
dX_t &= \mu_t^X \, dt + \sigma_t^X \, dB^Q_t \\
\mu_t^X &= K_0 + K_1 X_t \\
\sigma_t^X \sigma_t^{X'} &= H_0 + H_1 \cdot X_t.
\end{align*}
\]

and \( (K_0, K_1, H_0, H_1) \) satisfies the constraints in Duffie et al. (2003). In the spirit of DS, define two models to be equivalent if they gave the same prices for all securities with payoff depending on the short rate. Define two models to be affine equivalent if the models are related by \( Y = AX + b \) for some \( A, b \) with \( A \) nonsingular. Since

\[
\begin{align*}
\mu_t^Y &= AK_0 + AK_1 X_t \\
&= AK_0 - AK_1 A^{-1} b + AK_1 A^{-1} Y_t \\
\sigma_t^Y \sigma_t^{Y'} &= AH_0 A^t + AH_1 A^t X_1^t + \ldots + AH_1 A^t X_M^t \\
&= AH_0 A^t + \sum_{i=1}^M AH_1 A^t (A^{-1} b)_i + \\
&\sum_{i=1}^M AH_1 A^t (A^{-1})_{i1} Y_1^t + \ldots + \sum_{i=1}^M AH_1 A^t (A^{-1})_{iN} Y_N^t
\end{align*}
\]
Θₓ and Θᵧ are affine equivalent if there exists a A, b such that \( K₀^Y = AK₀^X - AK₁^X A^{-1} b, K₁^Y = AK₁^X A^{-1} \), etc. Finally, define a N-dimensional affine term structure model to be non-degenerate if it is not equivalent to an \( N - 1 \)-dimensional model. Theorem 1.4 gives conditions for an \( A_M(N) \) model to not degenerate to an \( A_M(N - 1) \) model. Similar eigenvector conditions and additional constraints are needed for an \( A_M(N) \) model to not degenerate to an \( A_{M-1}(N) \) model.

Equivalence and affine equivalence are closely related:

**Theorem A.1.** Any two equivalent non-degenerate DFS affine term structure models are affine equivalent

Theorem 1.1 then proves that the representation is unique.
B Non-diagonalizable Drift Feedback

A matrix is said to be in Jordan normal form if it is block diagonal with each block being a matrix which is constant along the diagonal with ones immediately above the diagonal.

\[
K = \begin{bmatrix}
J_1 & 0 & 0 & 0 \\
0 & J_2 & 0 & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & 0 & J_n
\end{bmatrix}, \quad J_i = \begin{bmatrix}
\lambda_i & 1 & 0 & 0 \\
0 & \lambda_i & 1 & 0 \\
0 & 0 & \ldots & 1 \\
0 & 0 & 0 & \lambda_i
\end{bmatrix}
\]

Once an ordering of the blocks is fixed, this similarity is unique. The diagonal entries of the Jordan blocks are the eigenvalues of the matrix and the size of the Jordan blocks are determined by the difference between the algebraic and geometric multiplicity of the eigenvalues. This rotation may result in a complex matrix. While conceptually this is not problematic, it may be convenient to rotate once more to a real valued matrix, such as an ordered real Schur decomposition.

The Riccati equations will be uncoupled across Jordan blocks. If there are Jordan blocks of size greater than one, the solutions will be an exponential polynomial of the form:

\[
B_j(\tau) = c_{-1} + c_0 e^{K_{j,j} t} + c_1 t e^{K_{j,j} t} + \ldots c_n t^n e^{K_{j,j} t}
\]  

(4)
C Detailed Example

Consider here the case of an $A_1(4)$ USV model where $K_G$ is diagonalizable with eigenvalues $a \neq 0, 2a, \text{and } 4a,$ and $B_1$ proportional to $B_2$: $B_1 \propto B_2$. It must be that $B_1 = \frac{\rho_1}{a}(1 - e^{a\tau})$ by the initial conditions. We must also have $\Sigma_{1,33}^G = 0$, or otherwise there will be an $e^{2a\tau}$ term which will violate Lemma 1.3. The condition that $\Sigma_{1,33}^G = 0$ implies the third Gaussian factor with the highest rate of mean reversion will have constant volatility. Positive semi-definiteness requires $\Sigma_{1,31}^G = \Sigma_{1,32}^G = \Sigma_{1,13}^G = \Sigma_{1,23}^G = 0$. Similarly, $\Sigma_{1,12}^G = \Sigma_{1,21}^G = 0$, or there will be a term $e^{3a\tau}$ which can’t cancel. This means the first two Gaussian factors must have constant covariance. Substituting into (2) gives an equation of the form:

$$0 = c_0 + c_1e^{at} + c_2e^{2at} + c_4e^{4at},$$

and four constraints $c_0 = c_1 = c_2 = c_4 = 0$.

One of the constraints is redundant, and finally there are three restrictions.

$$K_{2,1} = \frac{\rho_1 K_{2,2}^2 - \rho_1^2 - K_{1,1}\rho_1 K_{2,2} - \beta_{1,2}\rho_2^2}{\rho_2 K_{2,2}},$$

$$K_{3,1} = \frac{-\beta_{1,3}\rho_3^2 + 2\rho_1^2 + 2\beta_{1,2}\rho_2^2}{2\rho_1 K_{2,2}},$$

$$K_{4,1} = \frac{\beta_{1,3}\rho_3^2}{2\rho_4 K_{2,2}}.$$

Where $\beta_{1,i+1} = \Sigma_{1,i+1}^G$, which agrees with the notation in DS.

Another subcase will exist where $B_1$ is proportional to $B_3$. Additionally, there is a case with complex eigenvalues. Generally for $A_M(N)$ there will be many non-equivalent varieties of USV models.
D Proofs

Proof of Theorem 1.4. First note that the condition of having an eigenvector of $K_G$ orthogonal to $\rho_G$ is invariant to rotating the factors. That is, if we transform to $Y_t = AX_t$, then $K^Y_G = AK^X_G A^{-1}$ and $\rho^Y_G = (A^{-1})^\top \rho^X_G$. Hence if $x$ is an eigenvector of $K^X_G$ orthogonal to $\rho_G$, then $Ax$ is an eigenvector of $K^Y_G$ and is orthogonal to $\rho^Y_G$ since $(Ax) \cdot \rho_G = x^\top A^\top (A^\top)^{-1} \rho^X_G = x^\top \rho_G = 0$

To prove necessity, suppose there is such an eigenvector, $x$, of $K_G$. Consider first the case where the eigenvalue is real. We can extend to an orthonormal basis for $\mathbb{R}^{N-M}$: $\{x_{M+1}, \ldots x_{N-1}, x\}$. Rotating to this new set of coordinates:

$$K_G = \begin{bmatrix} K_a & 0 \\
K_b & k \end{bmatrix}, \quad \rho_N = 0.$$

$Z = (X_1, \ldots, X_{N-1})$ will have local conditional mean and variance which depend only on $Z$. This means $Z$ will be an $(N-1)$-dimensional Markov process. Since $r_t$ will depend only on $Z_t$, as $\rho_N = 0$, the model is in fact an $A_M(N-1)$ model.

If $x$ is a complex eigenvector orthogonal to $\rho_G$, then the complex conjugate will be an eigenvector also orthogonal to $\rho_G$. Similar to the real case, if we rotate to a real Schur decomposition with the eigenvectors orthogonal to $\rho_G$ in the last $2 \times 2$ block, we will find that the short rate is a function of only a (real-valued) $N-2$ dimensional state variable.

The sufficiency can be established by rotating $K_G$ into Jordan normal form (which may be complex). The ODEs for each Jordan block will be uncoupled. For a Jordan block of size $m$ there will be $m$ loadings of the form in (4) with $n = 0, 1, \ldots m - 1$. The coefficient of the highest order term depends only on the entry of $\rho_G$ for the last row of the Jordan block and will necessarily be non-zero by the assumption that no eigenvectors are orthogonal to $\rho_G$. 

29
Therefore, by Lemma 1.3 the loadings are linearly independent. It follows that for any rotation the loadings will be linearly independent (which means the fact that this rotation may be complex is not problematic). The fact that the loadings are independent precludes there existing any rotation where $B_i \equiv 0$ for some $i$ which is a necessary condition for the model to reduce to an $A_M(N - 1)$ model.

Proof of Theorem A.1. Suppose $\Theta_X$ is equivalent to $\Theta_Y$. Define $A^X_{T,s}, B^X_{T,s}$ by $E_t[e^{srT}] = e^{A^X_{T,s}+B^X_{T,s}X_t}$. By the non-degeneracy assumption, there must exist $\{(s_i, T_i)\}_{i=1}^N$ such that $B^X = [B^X_{s_1, T_1}, \ldots, B^X_{s_N, T_N}]$ has full rank. Otherwise, we can choose a lower dimensional set of linear combinations of the state variables which span $\{B^X_{s,T}\}$ and $r_t$ will be a function of this lower dimensional variable. By the fact that the lower dimensional variable can be inverted from prices, it can be shown to be Markov.

Since $E_t[e^{srT}] = e^{A^Y_{T,s}+B^Y_{T,s}X_t}$, it must be that $A^X + B^X X_t = A^Y + B^Y Y_t$, proving $\Theta_X$ and $\Theta_Y$ are affine equivalent.

Proof of Theorem 1.1. To see there is a unique element, suppose $\Theta_Y = \Theta_X$ and $Y = AX + b$. By considering the state space, $b_i = 0$ for $i \leq M$. Similarly, $A_{ij} = 0$ for $i \leq M, j > M$. The condition of non-correlated innovations between the CIR and Gaussian factors requires $A_{ij} = 0$ for $i > M, j \leq M$. The uniqueness of the ordered Jordan decomposition then implies the lower right $(N - M) \times (N - M)$ block of $A$ is the identity. In order that the state space $\mathbb{R}_+^M \times \mathbb{R}^{N-M}$ is preserved, the upper left $M \times M$ block of $A$ must be a permutation matrix or the identity. The assumed unique ordering of $\rho_1$ enforces it to be the identity. Thus, $A$ is the identity. $b_i = 0$ for $i > M$ by the normalization that $K_{ii} = 0$ for $i > M$ and the assumption that $K_{ii} \neq 0$ (no unit roots.) This verifies $\Theta_Y = \Theta_X$. 30
To prove existence, first begin with an affine process satisfying the conditions in Duffie et al. (2003).

1. **Remove local correlation between CIR and Gaussian factors**
   That is, for \( m \leq M, j \neq m \) set \( H_{1,mj}^m = H_{1,jm}^m = 0 \). To do this, transform by \( Y_t = AX_t \) where \( A \) is the identity except the \((j, m)\) entry which is \( \alpha \equiv -H_{1,mj}^m / H_{1,mm}^m \).

   For example, in the \( A_2(4) \) case, \( H_1 \) takes the form:

   \[
   H_1^1 = \begin{bmatrix}
   H_{1,11}^1 & 0 & H_{1,34}^1 & H_{1,14}^1 \\
   0 & 0 & 0 & 0 \\
   * & 0 & H_{1,33}^1 & H_{1,34}^1 \\
   * & 0 & * & H_{1,44}^1
   \end{bmatrix}, \quad H_1^2 = \begin{bmatrix}
   0 & 0 & 0 & 0 \\
   0 & H_{1,22}^2 & H_{1,23}^2 & H_{1,24}^2 \\
   0 & * & H_{1,33}^2 & H_{1,34}^2 \\
   0 & * & * & H_{1,44}^2
   \end{bmatrix},
   \]

   Where * are used for the symmetric entries in the matrices. In this case,

   \[
   A = \begin{bmatrix}
   1 & 0 & 0 & 0 \\
   0 & 1 & 0 & 0 \\
   -H_{1,31}^1 / H_{1,11} & 0 & 1 & 0 \\
   0 & 0 & 0 & 1
   \end{bmatrix},
   \]

   And the mapping \( Y = AX \) or \( X = A^{-1}Y \) is:

   \[
   \begin{align*}
   Y^1 &= X^1 & X^1 &= Y^1 \\
   Y^2 &= X^2 & X^2 &= Y^2 \\
   Y^3 &= X^3 + \alpha X_1 & X^3 &= Y^3 - \alpha Y_1 \\
   Y^4 &= X^4 & X^4 &= Y^4
   \end{align*}
   \]

   Multiplying on the left by \( A \) corresponds to subtracting \( H_{1,31}^1 / H_{1,11} \times \) (row 1) from (row 3). Multiplying on the right by \( A^t \) corresponds to
subtracting $H_{1,31}^2/H_{1,11}$ from (column 1)

The conditional covariance of $Y$ is

$$\sigma_t^Y \sigma_t^{Y'} = A H_0 A^t + A H_1^1 A^t X^1 + A H_2^2 A^t X^2$$

Since the first row and column of $H_0$ and $H_1^2$ are zero, $AH_0 A^t = H_0$, $AH_1^2 A^t = H_1^2$, and

$$\sigma_t^Y \sigma_t^{Y'} = H_0 + \begin{bmatrix} H_{1,11}^1 & 0 & 0 & H_{1,14}^1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{H}_{1,33}^1 & \hat{H}_{1,34}^1 \\ * & 0 & * & H_{1,44}^1 \end{bmatrix} X^1_t + H_1^2 X^2_t$$

Where the * denotes elements which have changed. Using that $Y^1 = X^1$ and $Y^2 = X^2$,

$$\sigma_t^Y \sigma_t^{Y'} = H_0 + \begin{bmatrix} H_{1,11}^1 & 0 & 0 & H_{1,14}^1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{H}_{1,33}^1 & \hat{H}_{1,34}^1 \\ * & 0 & * & H_{1,44}^1 \end{bmatrix} Y^1_t + H_1^2 Y^2_t$$

Which shows that this transformation has successfully removed the local correlation between the first and third factors.

$K_1$ still has the same form since multiplying on the left by $A$ adds a multiple of row 1 to row 3 and multiplying on the right by $A^{-1}$ adds a multiple of column 3 to column 1. Both of these operations preserve the structure of $K_1$ and $K_0$.

Repeated applications removes the local correlation between any CIR factor and any Gaussian factor. After these transformations, $H_1$ will be
in the form:

\[
H^1_1 = \begin{bmatrix}
H^1_{1,11} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & H^1_{1,33} & H^1_{1,34} \\
0 & 0 & * & H^1_{1,44}
\end{bmatrix}, \quad H^2_1 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & H^2_{1,22} & 0 & 0 \\
0 & 0 & H^2_{1,33} & H^2_{1,34} \\
0 & 0 & * & H^2_{1,44}
\end{bmatrix},
\]

2. **Rotate so that** $K_G$ **is in an ordered Jordan canonical form**

This is simply done by choosing $U$ so that $UKGU^{-1}$ is in the ordered Jordan normal form. $U$ can be formed by applying Gram-Schmidt to an ordered set of generalized eigenvectors. Transforming by a matrix $A$ which is block diagonal with the identity in the upper left $M \times M$-block and $U$ in the lower right block will have the desired effect.

3. **Order the CIR factors** This is done by transforming with a permutation matrix.

4. **Normalize the level of the Gaussian factors** This is done by a translation.

5. **Normalize the scale of the factors** This is done by a scaling transformation.

\[\square\]

*Proof of Theorem 1.2.* Suppose $\Theta = (K_0, K_1, \Sigma^G_0, \Sigma^G_1, \ldots, \Sigma^G_M)$ corresponds to a drift normalized model which is equivalent to a DS normalized model. By the same argument in the proof of Theorem A.1, the affine transformation relating the two must be block diagonal with the upper left block being the identity (assuming we enforce the permutation normalization in the DS version). Let $A$ give the lower right block of the affine transformation which
relates the Gaussian factors. Then $A^t \Sigma_0 G A = I$ and $A^t \Sigma_i G A$ are diagonal. From $A^t \Sigma_0 G A = I$, it follows that $\Sigma_0 G A = (A^{-1})^t$. Similarly $\Sigma_i V A = (A^{-1})^t D = \Sigma_0 G A D$, or $(\Sigma_0 G)^{-1} \Sigma_i G A = AD$ From this it follows that the each column of $A$ is an eigenvector of $(\Sigma_0 G)^{-1} \Sigma_i V$

Conversely, suppose the eigenvector condition holds. By first transforming by $A = \Sigma_0^{-1/2}$, it is enough to show that $S_i = \Sigma_0^{-1/2} \Sigma_i G \Sigma_0^{-1/2}$. The condition implies that $S_i y_j = \lambda_{i,j} y_j$ where $y_j = \Sigma^{1/2} x_j$. And so,

$$y_k^t S_i y_j = \lambda_{i,j} y_k^t y_j$$

$$= \lambda_{j,i} y_k^t y_j$$

Consider first the case where for any $j \neq k$ there exists an $i$ such that that $\lambda_{i,j} \neq \lambda_{i,k}$. Then by above, $y_k^t y_j = 0$ and so $y_k^t S_i y_j = 0$ for all $i$. This means $Y^t S_i Y$ is diagonal giving the result.

The other case is where there are $j \neq k$ such that $\lambda_{i,j} = \lambda_{i,k}$ for all $k$. For simplicity, suppose there is only one such pair. Let $V = \text{span}(y_j, y_k)$ and choose an orthonormal basis $\{\tilde{y}_j, \tilde{y}_k\}$. Then $S_i \tilde{y}_j = \lambda_{i,j} \tilde{y}_j$ and $S_i \tilde{y}_k = \lambda_{i,k} \tilde{y}_k$. Now using the same argument as before $\tilde{Y}^t S_i \tilde{Y}$ is diagonal where $\tilde{Y}$ is formed from $Y$ by replacing the columns $y_j$ and $y_k$ with $\tilde{y}_j, \tilde{y}_k$.

Geometrically, the Theorem 1.2 can be read interpreted as saying that for $N - M$ ellipsoids, if we first choose one $(\Sigma_0 G)$ to tranform to a sphere, we can simultaneously untill the remaining ellipsoids only if they have the same axes. (In the degenerate cases, some ellipsoids may have circular cross-sections which gives flexibility in choice of axes.) The proof is a standard application of the theory of generalized eigenvalues.

Similar results where the eigenvector condition is replaced with generalized eigenvector conditions apply in the case that $\Sigma_0$ is merely positive semi-
definite and not strictly positive definite (which the DS normalization precludes.)
References


Notes

1 See Dai and Singleton (2000) for a description of affine term structure models. They classify affine term structure models into non-nested families denoted $A_M(N)$. $N$ is the total number of factors and $M$ is the number of factors driving volatility.

2 If there are $D$ locally deterministic factors (i.e. $\dim(\{H_1 \cdot x\}_{x \in \mathbb{R}^N} = N - D)$), there are incomplete markets if the projection of $\{B(\tau)\}_{\tau \in \mathbb{R}^+}$ onto the stochastic factors has dimension less than $N - D$.

3 The interpretation that it is volatility which is unspanned relies on using yields for the other factors ($Y^2, Y^3$). If there are other observable factors, different rotations are possible where volatility has a non-zero loading.

4 More generally, $K_G$ can be in ordered Jordan normal form, but for ease of exposition I focus on the diagonal case in the main text. Appendix B gives details for the more general Jordan form. However, there are economically interesting cases where $K_G$ is non-diagonalizable. For instance, Theorem 1.4 below implies that the more general Jordan form is necessary in order to have two unit roots driving the short rate.

5 Another possibility suggested by these facts is to consider a jump diffusion model. The present analysis is restricted to diffusive affine term structure models.

6 There are also cases with complex eigenvalues where there will be $2 \times 2$ blocks on the diagonal or with with real eigenvalues and Jordan blocks of size greater than 2. For simplicity, these and other special subcases, such as unit roots where the level cannot be normalized or locally deterministic CIR-type factors, are not considered. They do, however, represent other canonical forms.
Figure 1: Zero Coupon Yields
This figure plots 6 month, 2 year, 5 year, and 10-year swap implied zero coupon yields for the period April 1987 to September 2005.
This figure plots EGARCH(1,1) estimates of the conditional variance of 6 month, 2 year, 5 year, and 10-year swap implied zeros for the period April 1987 to September 2005.
This figure plots the factor loadings for zero coupon yields, $B_\nu(\tau) \equiv -B(\tau), \tau = (1 - e^{k\tau})/(k\tau)$. The factor loadings show how the yield curve shifts across maturities due to shocks in a Gaussian risk factor driving the short rate with a given level of mean reversion. The values of $k$ are chosen so that the half-life of the factor, $-\log(2)/k$, corresponds to 3 months, 1 year, 5 years, and 50 years.
Figure 4: Yield variance principal component loadings
This figure plots the first principal component loadings of changes in
EGARCH(1,1) estimates of the conditional variance of 3 month, 6 month, 1-, 
2-, 3-, 4-, 5-, 6-, 7-, 8-, 9- and 10-year swap implied zeros for the period April
1987 to September 2005. The first principal component explained 67.5% of
the variance.
This figure plots the cross sectional changes in conditional variances when the cross section changes according to

$$c_1 \frac{(1 - e^{-k_1 \tau})^2}{(k_1 \tau)^2} + c_2 \frac{(1 - e^{-k_2 \tau})^2}{(k_2 \tau)^2}$$

using various weights and mean reversions corresponding to either 50 years and 1 year or 5 years and 1 year.

Figure 5: Variance Loadings