Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields?
An Investigation of Unspanned Stochastic Volatility

Pierre Collin-Dufresne
Robert S. Goldstein
Christopher S. Jones

VSAM, July 2007
Term structure models in theory and reality

- In theory, no-arbitrage multifactor models should prove useful in many contexts, including:
  - Asset allocation
  - Asset/liability management
  - Consistent pricing/hedging of caps and swaptions
  - Optimal exercise policy
Term structure models in theory and reality

• In theory, no-arbitrage multifactor models should prove useful in many contexts, including:
  – Asset allocation
  – Asset/liability management
  – Consistent pricing/hedging of caps and swaptions
  – Optimal exercise policy

• In reality:
  – Portfolio managers use PC analysis
  – Option traders use one-factor models (e.g. HW, BDT)
Term structure models in theory and reality

• In theory, no-arbitrage multifactor models should prove useful in many contexts, including:
  – Asset allocation
  – Asset/liability management
  – Consistent pricing/hedging of caps and swaptions
  – Optimal exercise policy

• In reality:
  – Portfolio managers use PC analysis
  – Option traders use one-factor models (e.g. HW, BDT)

• Few use dynamic term structure (affine) models!
  – High cost: hard to estimate and implement
  – Low benefit: unsatisfactory out-of sample performance for forecasting, pricing and hedging
Term structure models and yield dynamics

• Litterman and Scheinkman (1991): 3 factors are necessary to explain yield dynamics
• Engle, Lilien, and Robins (1987): yield volatility is stochastic
• Litterman, Scheinkman and Weiss (1991): curvature factor related to interest rate volatility
• Dai and Singleton (2000): the 3-factor affine “$A_1(3)$” model seems to work best at capturing both
Term structure models and yield dynamics

- Litterman and Scheinkman (1991): 3 factors are necessary to explain yield dynamics
- Engle, Lilien, and Robins (1987): yield volatility is stochastic
- Litterman, Scheinkman and Weiss (1991): curvature factor related to interest rate volatility
- Dai and Singleton (2000): the 3-factor affine “A₁(3)” model seems to work best at capturing both

- Some strong results from the affine class: if $X_t$ is the state vector, then in general volatility is spanned:
  
  \[ Y_t = a_0 + a^\top X_t \]
  
  \[ \text{Var}(dY_t) = b_0 + b^\top X_t \]
  
  \[ \Rightarrow \text{Var}(dY_t) = c_0 + c^\top Y_t \]

- If volatility is spanned by the yield curve, then what is the role for the derivatives market?
Term structure models and yield dynamics

- Litterman and Scheinkman (1991): 3 factors are necessary to explain yield dynamics
- Engle, Lilien, and Robins (1987): yield volatility is stochastic
- Litterman, Scheinkman and Weiss (1991): curvature factor related to interest rate volatility
- Dai and Singleton (2000): the 3-factor affine “A₁(3)” model seems to work best at capturing both

- Some strong results from the affine class: if \( X_t \) is the state vector, then in general volatility is spanned:
  \[- Y_t = a_0 + a^\top X_t \]
  \[- \text{Var}(dY_t) = b_0 + b^\top X_t \]
  \[\Rightarrow \text{Var}(dY_t) = c_0 + c^\top Y_t \]

- If volatility is spanned by the yield curve, then what is the role for the derivatives market?
- Collin-Dufresne and Goldstein (2002): under certain conditions, volatility will be unspanned, e.g.,
  \[ Y_t = a_0 + a^\top X_t \quad \text{BUT} \quad \text{Var}(dY_t) \neq c_0 + c^\top Y_t \]
Term structure models and yield dynamics

- Litterman and Scheinkman (1991): 3 factors are necessary to explain yield dynamics
- Engle, Lilien, and Robins (1987): yield volatility is stochastic
- Litterman, Scheinkman and Weiss (1991): curvature factor related to interest rate volatility
- Dai and Singleton (2000): the 3-factor “$A_1(3)$” model seems to work best at capturing both

- Some strong results from the affine class: if $X_t$ is the state vector, then in general volatility is spanned:
  \[ Y_t = a_0 + a^\top X_t \]
  \[ \text{Var}(dY_t) = b_0 + b^\top X_t \]
  \[ \Rightarrow \text{Var}(dY_t) = c_0 + c^\top Y_t \]

- If volatility is spanned by the yield curve, then what is the role for the derivatives market?
- Collin-Dufresne and Goldstein (2002): under certain conditions, volatility will be unspanned, e.g.,
  \[ Y_t = a_0 + a^\top X_t \quad \text{BUT} \quad \text{Var}(dY_t) \neq c_0 + c^\top Y_t \]

- Can models in which volatility is spanned fit both the cross section of yields and the time series of volatilities?
- Can unspanned stochastic volatility models do better?
What is ‘Unspanned Stochastic Volatility’?

• Typical stochastic volatility term structure model: Longstaff and Schwartz (1992)
  
  \[ \begin{align*}
  dr &= \kappa_r(\theta_r - r)\,dt + \sqrt{V}\,dz_1^Q \\
  dV &= \kappa_V(\theta_V - V)\,dt + \sigma\sqrt{V}\,dz_2^Q
  \end{align*} \]

  
  • In this model zero-coupon bond prices given by:
  \[
  P_T(t) = e^{A(T-t)+B(T-t)r_t+C(T-t)V_t}
  \]

  ⇒ Volatility risk can be hedged with appropriate position in any two bonds.

  ⇒ Volatility plays dual role: cross-section vs. time-series.

Fong and Vasicek (91), Longstaff and Schwartz (92), Chen and Scott (93), Balduzzi et al. (96), Chen (96), DS (00) . . .
What is ‘Unspanned Stochastic Volatility’?

• Typical stochastic volatility term structure model: Longstaff and Schwartz (1992)

\[
\begin{align*}
    dr &= \kappa_r(\theta_r - r) \, dt + \sqrt{V} \, dz_1^Q \\
    dV &= \kappa_V(\theta_V - V) \, dt + \sigma\sqrt{V} \, dz_2^Q
\end{align*}
\]

• In this model zero-coupon bond prices given by:

\[
P_T(t) = e^{A(T-t) + B(T-t) r_t + C(T-t) V_t}
\]

⇒ Volatility risk can be hedged with appropriate position in any two bonds.

⇒ Volatility plays dual role: cross-section vs. time-series.

Fong and Vasicek (91), Longstaff and Schwartz (92), Chen and Scott (93), Balduzzi et al. (96), Chen (96), DS (00)...

• Contrasts with equity derivatives models such as Heston (1993):

\[
\begin{align*}
    \frac{dS}{S} &= r \, dt + \sqrt{V} \, dz_1^Q \\
    dV &= \kappa(\theta - V) \, dt + \sigma\sqrt{V} \, dz_2^Q
\end{align*}
\]

⇒ Volatility risk cannot be hedged by any portfolio of stock and bond

• Difference: equity models a \textit{traded} asset, and \textit{its} volatility.
**What is ‘Unspanned Stochastic Volatility’?**

- Typical stochastic volatility term structure model: Longstaff and Schwartz (1992)

\[ dr = \kappa_r(\theta_r - r) \, dt + \sqrt{V} \, dz^Q_1 \]

\[ dV = \kappa_V(\theta_V - V) \, dt + \sigma \sqrt{V} \, dz^Q_2 \]

- In this model zero-coupon bond prices given by:

\[ P^T(t) = e^{A(T-t)+B(T-t)r_t+C(T-t)V_t} \]

⇒ Volatility risk can be hedged with appropriate position in any two bonds.

⇒ Volatility plays dual role: cross-section vs. time-series.

Fong and Vasicek (91), Longstaff and Schwartz (92), Chen and Scott (93), Balduzzi et al. (96), Chen (96), DS (00)...

- Contrasts with equity derivatives models such as Heston (1993):

\[ \frac{dS}{S} = r \, dt + \sqrt{V} \, dz^Q_1 \]

\[ dV = \kappa(\theta - V) \, dt + \sigma \sqrt{V} \, dz^Q_2 \]

⇒ Volatility risk cannot be hedged by any portfolio of stock and bond

- Difference: Have modeled a **traded** asset, and **its** volatility.

- The idea of USV is to obtain the same relation for bond prices:

\[ \frac{dP^T(t)}{P^T(t)} = r_t \, dt - \sigma_P(T - t) \sqrt{V_t} \, dz^Q_1(t) \]

Andreasen et al.(97), Collin-Dufresne and Goldstein (02), Kimmel (03), Casassus, Collin-Dufresne, Goldstein (05), Trolle-Schwartz (07).
Outline of the talk

- New representation of affine models
- Introduce old and new USV models
- Describe estimation method
- Show some results
Affine models in “traditional” SDE form

- The $Q$ measure dynamics for **latent variables**

$$
dX(t) = \mathcal{K}^Q (\theta^Q - X(t)) \ dt + \Sigma \sqrt{S(t)} dZ^Q(t),
$$

where $\mathcal{K}^Q$ and $\Sigma$ are $(N \times N)$ and $S$ is diagonal with $S_{ii}(t) = \alpha_i + \beta_i^\top X(t)$

- The short rate $r(t) = \delta_0 + \delta_1^\top X(t)$, which implies that yields are affine in $X$:

$$
Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t),
$$

where $A(\tau)$ and $B(\tau)$ satisfy a system of ODEs.
Affine models in “traditional” SDE form

- The $Q$ measure dynamics for latent variables
  \[
  dX(t) = \mathcal{K}^Q \left( \theta^Q - X(t) \right) dt + \Sigma \sqrt{S(t)} dZ^Q(t),
  \]
  where $\mathcal{K}^Q$ and $\Sigma$ are $(N \times N)$ and $S$ is diagonal with $S_{ii}(t) = \alpha_i + \beta_i^\top X(t)$

- The short rate $r(t) = \delta_0 + \delta_1^\top X(t)$, which implies that yields are affine in $X$:
  \[
  Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t),
  \]
  where $A(\tau)$ and $B(\tau)$ satisfy a system of ODEs

- In the so-called “$A_m(N)$” model,
  - there are $N$ state variables
  - of which $m$ are square root processes that are the sole driver of stochastic volatility.

- Popular models:
  - $A_0(3)$: three state variables with constant covariances – a Gaussian model
  - $A_1(3)$: two “Gaussian” variables and one CIR process that drives stochastic volatility
Affine models in “traditional” SDE form

- The $Q$ measure dynamics for **latent variables**

  \[ dX(t) = \mathcal{K}^Q (\theta^Q - X(t)) \, dt + \Sigma \sqrt{S(t)} \, dZ^Q(t), \]

  where $\mathcal{K}^Q$ and $\Sigma$ are $(N \times N)$ and $S$ is diagonal with $S_{ii}(t) = \alpha_i + \beta_i^\top X(t)$

- The short rate $r(t) = \delta_0 + \delta_1^\top X(t)$, which implies that yields are affine in $X$:

  \[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t), \]

  where $A(\tau)$ and $B(\tau)$ satisfy a system of ODEs

- In the so-called “$A_m(N)$” model,
  - there are $N$ state variables
  - of which $m$ are square root processes that are the sole driver of stochastic volatility.

- Popular models:
  - $A_0(3)$: three state variables with constant covariances – a Gaussian model
  - $A_1(3)$: two Gaussian variables and one CIR process that drives stochastic volatility

- Several challenges:
  - There are far too many unknown parameters to estimate
  - Model is *inadmissible* in this general form
  - Variable are latent, i.e., devoid of economic meaning.

$\Rightarrow$ CGJ (2006): rewrite model in terms of “observable” state variables
Observable state variables and the mean/covariance form

- Instead of using standard latent variables, we write all the models in terms of
  - $r$: instantaneous short rate
  - $\mu^Q$: the Q drift of $r$
  - $\theta^Q$: the Q drift of $\mu^Q$
  - $V$: the instantaneous variance of $r$
Observable state variables and the mean/covariance form

- Instead of using standard latent variables, we write all the models in terms of
  - $r$: instantaneous short rate
  - $\mu^Q$: the Q drift of $r$
  - $\theta^Q$: the Q drift of $\mu^Q$
  - $V$: the instantaneous variance of $r$

- CGJ (2006) show that all are affine in the original state vector and all are observable independently of a model (and parameters).
  - $r(t) =$ level of the term structure for very short maturity.
  - $\mu^Q(t) =$ twice the slope of term structure for very short maturity.
  - $\theta^Q(t) =$ thrice the curvature of term structure for very short maturity plus $V(t)$. 
Observable state variables and the mean/covariance form

- Instead of using standard latent variables, we write all the models in terms of
  - \( r \): instantaneous short rate
  - \( \mu^Q \): the Q drift of \( r \)
  - \( \theta^Q \): the Q drift of \( \mu^Q \)
  - \( V \): the instantaneous variance of \( r \)

- CGJ (2006) show that all are affine in the original state vector and all are observable independently of a model (and parameters).
  - \( r(t) \) = level of the term structure for very short maturity.
  - \( \mu^Q(t) \) = twice the slope of term structure for very short maturity.
  - \( \theta^Q(t) \) = thrice the curvature of term structure for very short maturity plus \( V(t) \).

- In addition, all models are written in terms of instantaneous means and covariances:
  \[
  \frac{d}{dt} \mathbb{E}[dX_t] = a^Q + b^Q X_t
  \]
  \[
  \frac{d}{dt} \text{Cov}(dX_t, dX_t^\top) = \Omega_0 + \sum_{i=1}^{m} \Omega_i X_{i,t}
  \]

where \( X \) is some combination of \( r, \mu^Q, \theta^Q \), and/or \( V \).
Example: the $A_1(3)$ model

If we define the state vector as

$$X_t = \begin{bmatrix} r_t \\ \mu_t^Q \\ \mu_t \\ V_t \end{bmatrix}$$

then

$$\frac{1}{dt} \mathbb{E}^{Q} [dX_t] = \begin{bmatrix} \mu_t^Q \\ m_0 + m_r r_t + m_{\mu} \mu_t^Q + m_{\gamma} V_t \\ \gamma_V - \kappa_V V_t \end{bmatrix}$$

and

$$\frac{1}{dt} \text{Cov} \left( dX_t, dX_t^\top \right) = \begin{bmatrix} V & c_{r\mu} & 0 \\ c_{r\mu} & \sigma_{\mu} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & c_{r\mu} & c_{r\sigma} \\ c_{r\mu} & \sigma_{\mu} & c_{\mu\sigma} \\ c_{r\sigma} & c_{\mu\sigma} & \sigma_{\sigma} \end{bmatrix} (V_t - V)$$

- Each restriction is required by definition or for admissibility
- Level, slope, and volatility factors
- 24 parameters (14 risk neutral and 10 risk premia)
USV in the $A_1(3)$ model

Yields in the $A_1(3)$ model are of the form

$$Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B_r(\tau)}{\tau} r(t) + \frac{B_\mu(\tau)}{\tau} \mu^Q(t) + \frac{B_V(\tau)}{\tau} V(t)$$
USV in the $A_1(3)$ model

Yields in the $A_1(3)$ model are of the form

$$Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B_r(\tau)}{\tau} r(t) + \frac{B_\mu(\tau)}{\tau} \mu^Q(t) + \frac{B_V(\tau)}{\tau} V(t)$$

Unspanned stochastic volatility is a set of conditions that ensure $B_V(\tau) \equiv 0$. They are:

$$m_r = -2c_{r\mu}^2$$
$$m_\mu = 3c_{r\mu}$$
$$m_V = 1$$
$$\sigma_\mu = c_{r\mu}^2$$
USV in the $A_1(3)$ model

Yields in the $A_1(3)$ model are of the form

$$Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B_r(\tau)}{\tau} r(t) + \frac{B_\mu(\tau)}{\tau} \mu^Q(t) + \frac{B_V(\tau)}{\tau} V(t)$$

Unspanned stochastic volatility is a set of conditions that ensure $B_V(\tau) \equiv 0$. They are:

- $m_r = -2c_{r\mu}^2$
- $m_\mu = 3c_{r\mu}$
- $m_V = 1$
- $\sigma_\mu = c_{r\mu}^2$

- Volatility affects long term yields via two channels:
  - Direct impact on the drift of $\mu^Q$
  - Jensen’s inequality in

$$E^Q \left[ \exp \left( -\int_0^T r_t \, dt \right) \right]$$
**USV in the $A_1(3)$ model**

Yields in the $A_1(3)$ model are of the form

$$Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B_r(\tau)}{\tau} r(t) + \frac{B_\mu(\tau)}{\tau} \mu^Q(t) + \frac{B_V(\tau)}{\tau} V(t)$$

Unspanned stochastic volatility is a set of conditions that ensure $B_V(\tau) \equiv 0$. They are:

- $m_r = -2c_{r\mu}^2$
- $m_\mu = 3c_{r\mu}$
- $m_V = 1$
- $\sigma_\mu = c_{r\mu}^2$

- Volatility affects long term yields via two channels:
  - Direct impact on the drift of $\mu^Q$
  - Jensen’s inequality in

$$E^Q \left[ \exp \left( -\int_0^T r_t \, dt \right) \right]$$

- USV restrictions insure net effect cancels out
- Number of free parameters reduced by 4
- Parameters of $Q$ drift of $V(t)$ not identified from bond prices alone
Alternative: the $A_0(2)$ model

$$X_t = \begin{bmatrix} r_t \\ \mu_t^Q \end{bmatrix}$$

$$\frac{1}{dt} E^Q [dX_t] = \begin{bmatrix} \mu_t^Q \\ m_0 + m_r r_t + m_\mu \mu_t^Q \end{bmatrix}$$

$$\frac{1}{dt} \text{Cov} (dX_t, dX_t^\top) = \begin{bmatrix} V & \sigma_r \\ \sigma_r & \sigma_\mu \end{bmatrix}$$

- Level and slope factors
- Constant covariance matrix
- 12 parameters total
Alternative: the $A_1(4)$ model

$$X_t = \begin{bmatrix} r_t \\ \mu_t^Q \\ \theta_t^Q \\ V_t \end{bmatrix}$$

$$\frac{1}{dt}E^Q[dX_t] = \begin{bmatrix} 
\mu_t^Q \\
\theta_t^Q \\
a_0 + a_r r_t + a_\mu r_t Q_t + a_\theta \theta_t Q_t + a_V V_t \\
\gamma_V - \kappa_V V_t 
\end{bmatrix}$$

$$\frac{1}{dt} \text{Cov} \left( dX_t, dX_t^T \right) = \begin{bmatrix} 
V & c_{r\mu} & c_{r\theta} & 0 \\
c_{r\mu} & \sigma_\mu & c_{\mu\theta} & 0 \\
c_{r\theta} & c_{\mu\theta} & \sigma_\theta & 0 \\
0 & 0 & 0 & 0 
\end{bmatrix} + \begin{bmatrix} 
1 & c_{r\mu} & c_{r\theta} & c_{rV} \\
c_{r\mu} & \sigma_\mu & c_{\mu\theta} & c_{\muV} \\
c_{r\theta} & c_{\mu\theta} & \sigma_\theta & c_{\thetaV} \\
c_{rV} & c_{\muV} & c_{\thetaV} & \sigma_V 
\end{bmatrix} (V_t - \bar{V})$$

- Level, slope, curvature, and volatility are now all distinct.
- USV conditions impose 6 constraints, resulting in 13 Q parameters (fewer than unrestricted $A_1(3)$) and 17 risk premia parameters.
- $\gamma_V$ and $\kappa_V$ are still not identified from bond prices under USV.
Econometric approach

Most econometric approaches are based on the idea that

\[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t) \]

is an invertible transformation. Two common methods:
Most econometric approaches are based on the idea that

\[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau}X(t) \]

is an invertible transformation. Two common methods:

- The “inversion” approach
  1. In a 3-factor model, assume there are 3 yields observed “without error”
  2. Invert those 3 yields to get 3 factors
  3. Use observed factors to compute likelihood: \( p(Y_t | X_{t-1}) \)
Most econometric approaches are based on the idea that

\[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau}X(t) \]

is an invertible transformation. Two common methods:

- **The “inversion” approach**
  1. In a 3-factor model, assume there are 3 yields observed “without error”
  2. Invert those 3 yields to get 3 factors
  3. Use observed factors to compute likelihood: \( p(Y_t | X_{t-1}) \)

- **Put factors and yields in state space form and use Kalman filter.**
  1. Autoregressive state equation: \( X_t = a + bX_{t-1} + \epsilon_t \)
  2. Linear observation equation: \( Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau}X(t) + \eta_t \)
Econometric approach

Most econometric approaches are based on the idea that

\[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t) \]

is an invertible transformation. Two common methods:

- The “inversion” approach
  1. In a 3-factor model, assume there are 3 yields observed “without error”
  2. Invert those 3 yields to get 3 factors
  3. Use observed factors to compute likelihood: \( p(Y_t|X_{t-1}) \)

- Put factors and yields in state space form and use Kalman filter.
  1. Autoregressive state equation: \( X_t = a + bX_{t-1} + \epsilon_t \)
  2. Linear observation equation: \( Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t) + \eta_t \)

Problem: neither one works when some elements of \( B(\tau) \) are zero.
Econometric approach

Most econometric approaches are based on the idea that

\[ Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t) \]

is an invertible transformation. Two common methods:

- The “inversion” approach
  1. In a 3-factor model, assume there are 3 yields observed “without error”
  2. Invert those 3 yields to get 3 factors
  3. Use observed factors to compute likelihood: \( p(Y_t|X_{t-1}) \)

- Put factors and yields in state space form and use Kalman filter.
  1. Autoregressive state equation: \( X_t = a + bX_{t-1} + \epsilon_t \)
  2. Linear observation equation: \( Y(t, \tau) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X(t) + \eta_t \)

Problem: neither one works when some elements of \( B(\tau) \) are zero.

Solution: MCMC
Econometric approach

• Bayesian MCMC related to Jones (2003), Lamoureaux and Witte (2002), Bester (2004), Sanford and Martin (2003), and Polson, Stroud, and Müller (2001)

• Diffusion process:

\[ dX_t \sim N \left( \left( a + bX_t \right) dt, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) dt \right) \]

discretized using the Euler approximation.

\[ X_{t+h} - X_t \sim N \left( \left( a + bX_t \right) h, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) h \right) \]
Econometric approach

• Bayesian MCMC related to Jones (2003), Lamoureux and Witte (2002), Bester (2004), Sanford and Martin (2003), and Polson, Stroud, and Müller (2001)

• Diffusion process:
\[
\begin{align*}
    dX_t &\sim N \left( \left( a + bX_t \right) dt, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) dt \right) \\
\end{align*}
\]
discretized using the Euler approximation.

\[
\begin{align*}
    X_{t+h} - X_t &\sim N \left( \left( a + bX_t \right) h, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) h \right) \\
\end{align*}
\]

• Principal components are used in place of yields and are assumed measured with error:
\[
\mathcal{P}_t = K + LX_t + e_t,
\]
where \(e_t\) is Gaussian.
Econometric approach

- Bayesian MCMC related to Jones (2003), Lamoureux and Witte (2002), Bester (2004), Sanford and Martin (2003), and Polson, Stroud, and Müller (2001)

- Diffusion process:

  \[ dX_t \sim N \left( (a + bX_t) \, dt, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) \, dt \right) \]

  discretized using the Euler approximation.

  \[ X_{t+h} - X_t \sim N \left( (a + bX_t) \, h, \left( \Omega_0 + \Omega_V (V_t - \bar{V}) \right) \, h \right) \]

- Principal components are used in place of yields and are assumed measured with error:

  \[ P_t = K + LX_t + e_t, \]

  where \( e_t \) is Gaussian.

- Observed yield data augmented with unobserved factor data.
  
  - Unobserved \( V_t \) drawn for each \( t \) separately using Metropolis-Hastings.
  
  - Other state variables drawn simultaneously for all \( t \) using the simulation smoother.
Econometric approach

- Bayesian MCMC related to Jones (2003), Lamoureaux and Witte (2002), Bester (2004), Sanford and Martin (2003), and Polson, Stroud, and Müller (2001)

- Diffusion process:
  \[ dX_t \sim N \left( \left( a + bX_t \right) dt, \left( \Omega_0 + \Omega_V(V_t - \overline{V}) \right) dt \right) \]
  discretized using the Euler approximation.
  \[ X_{t+h} - X_t \sim N \left( \left( a + bX_t \right) h, \left( \Omega_0 + \Omega_V(V_t - \overline{V}) \right) h \right) \]

- Principal components are used in place of yields and are assumed measured with error:
  \[ P_t = K + LX_t + e_t, \]
  where \( e_t \) is Gaussian.

- Observed yield data augmented with unobserved factor data.
  - Unobserved \( V_t \) drawn for each \( t \) separately using Metropolis-Hastings.
  - Other state variables drawn simultaneously for all \( t \) using the simulation smoother.

- Low frequency data augmented with high frequency data.
  - Does not seem to make a difference.
Data


- 2003-2005 is a hold-out sample.

- Coupon yields are bootstrapped to estimate zero coupon yields (.5, 1, 2, 3, 4, 5, 7, and 10 years).

- Principal components are computed from yields to try to orthogonalize measurement errors.
  - Yields are linear in factors, and principal components are linear in yields.
  - There exist $K$ and $L$ such that $\mathcal{P}_t = K + LX_t$
Results outline

- Parameter estimates
- Specification analysis
- Yield curve fit
- Forecasting yield volatilities
- Interpreting $\hat{V}_t$
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$A_i(2)$</th>
<th>$A_i(3)$ USV</th>
<th>$A_i(3)$</th>
<th>$A_i(4)$ USV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_0$</td>
<td>0.003</td>
<td>0.003</td>
<td>0.211</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002, 0.003)</td>
<td>(0.002, 0.003)</td>
<td>(0.188, 0.238)</td>
<td></td>
</tr>
<tr>
<td>$m_r$</td>
<td>-0.019</td>
<td>-0.031</td>
<td>-1.238</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.025, -0.014)</td>
<td>(-0.032, -0.028)</td>
<td>(-1.348, -1.131)</td>
<td></td>
</tr>
<tr>
<td>$m_\mu$</td>
<td>-0.633</td>
<td>-0.371</td>
<td>-2.804</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.705, -0.585)</td>
<td>(-0.382, -0.354)</td>
<td>(-2.959, -2.597)</td>
<td></td>
</tr>
<tr>
<td>$m_V$</td>
<td>1.000</td>
<td></td>
<td></td>
<td>-923.149</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-1003.107, -811.230)</td>
</tr>
<tr>
<td>$\gamma_P \times 10^4$</td>
<td>1.015</td>
<td>0.568</td>
<td>0.803</td>
<td></td>
</tr>
<tr>
<td>$\kappa_P$</td>
<td>1.130</td>
<td>0.441</td>
<td>1.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.537, 0.484)</td>
<td>(0.756, 0.056)</td>
<td>(2.116, 0.412)</td>
<td></td>
</tr>
<tr>
<td>$V \times 10^4$</td>
<td>1.021</td>
<td>0.020</td>
<td>0.111</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.883, 1.193)</td>
<td>(0.003, 0.041)</td>
<td>(0.046, 0.298)</td>
<td>(0.000, 0.011)</td>
</tr>
<tr>
<td>$\sigma_\mu \times 10^4$</td>
<td>0.981</td>
<td>0.182</td>
<td>5.809</td>
<td>3.448</td>
</tr>
<tr>
<td></td>
<td>(0.643, 1.497)</td>
<td>(0.158, 0.207)</td>
<td>(3.208, 11.349)</td>
<td>(2.675, 3.999)</td>
</tr>
<tr>
<td>$c_{r\mu} \times 10^4$</td>
<td>-0.376</td>
<td>0.043</td>
<td>0.708</td>
<td>-0.102</td>
</tr>
<tr>
<td></td>
<td>(-0.637, -0.240)</td>
<td>(0.017, 0.071)</td>
<td>(0.187, 0.926)</td>
<td>(-0.197, -0.014)</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.015</td>
<td></td>
<td>17.660</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.014, 0.016)</td>
<td></td>
<td>(11.327, 24.871)</td>
<td>(0.007, 0.008)</td>
</tr>
<tr>
<td>$\sigma_V \times 10^4$</td>
<td>1.647</td>
<td>0.025</td>
<td>1.221</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.948, 2.845)</td>
<td>(0.022, 0.030)</td>
<td>(0.811, 1.968)</td>
<td></td>
</tr>
<tr>
<td>$c_{rV} \times 10^3$</td>
<td>-0.124</td>
<td>-3.617</td>
<td>-0.087</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.127, -0.118)</td>
<td>(-4.526, -2.893)</td>
<td>(-0.091, -0.083)</td>
<td></td>
</tr>
<tr>
<td>$c_{\mu V}$</td>
<td>-0.003</td>
<td>-0.376</td>
<td>2.685</td>
<td>-0.233</td>
</tr>
<tr>
<td></td>
<td>(-2.430, 3.002)</td>
<td>(-0.504, -0.209)</td>
<td>(0.152, 4.457)</td>
<td>(-0.233, -0.233)</td>
</tr>
</tbody>
</table>
Specification analysis

In the Euler approximation

\[ X_{t+h} - X_t \sim N \left( \left( a + bX_t \right) h, \left( \Omega_0 + \Omega_V(V_t - V) \right) h \right) \]

the standardized residuals should be i.i.d. N(0,1). Is this the case?

Table 3: Specification analysis

<table>
<thead>
<tr>
<th></th>
<th>(A_0(2))</th>
<th>(A_1(3)) USV</th>
<th>(A_1(3))</th>
<th>(A_1(4)) USV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(r) (\mu^2)</td>
<td>(r) (\mu^2)</td>
<td>(V)</td>
<td>(r) (\mu^2)</td>
</tr>
<tr>
<td>Mean (\hat{\epsilon})</td>
<td>0.00 0.00</td>
<td>0.01 0.00</td>
<td>0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>SD of (\hat{\epsilon})</td>
<td>0.97 0.99</td>
<td>1.00 1.12</td>
<td>0.99*</td>
<td>0.95 0.96*</td>
</tr>
<tr>
<td>Skewness of (\hat{\epsilon})</td>
<td>-0.80* 0.31</td>
<td>-0.03 0.04</td>
<td>0.06</td>
<td>-0.35 0.03*</td>
</tr>
<tr>
<td>Kurtosis of (\hat{\epsilon})</td>
<td>7.42* 3.93</td>
<td>3.62 4.13*</td>
<td>2.96*</td>
<td>7.21* 4.04*</td>
</tr>
<tr>
<td>1st order AC of (\hat{\epsilon})</td>
<td>0.19* 0.24</td>
<td>0.10 -0.02*</td>
<td>-0.01</td>
<td>0.05* 0.04*</td>
</tr>
<tr>
<td>1st order AC of (</td>
<td>\hat{\epsilon}</td>
<td>)</td>
<td>0.14* 0.07</td>
<td>0.00 -0.01</td>
</tr>
<tr>
<td>1st order AC of (\hat{\epsilon}^2)</td>
<td>0.05* 0.04</td>
<td>-0.01 -0.01</td>
<td>0.00</td>
<td>0.16* 0.08*</td>
</tr>
</tbody>
</table>
Table 4: In-sample yield fits

<table>
<thead>
<tr>
<th></th>
<th>$A_0(2)$</th>
<th>$A_1(3)$ USV</th>
<th>$A_1(3)$</th>
<th>$A_1(4)$ USV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mean $\hat{e}$ (basis points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>0.05</td>
<td>-5.38**</td>
<td>-0.11</td>
<td>0.69</td>
</tr>
<tr>
<td>1-year</td>
<td>-0.15</td>
<td>-1.13</td>
<td>0.14</td>
<td>-0.70</td>
</tr>
<tr>
<td>2-year</td>
<td>-0.06</td>
<td>3.25**</td>
<td>0.06</td>
<td>-0.79</td>
</tr>
<tr>
<td>3-year</td>
<td>0.14</td>
<td>4.13**</td>
<td>-0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>4-year</td>
<td>0.20</td>
<td>3.13**</td>
<td>-0.08</td>
<td>0.69</td>
</tr>
<tr>
<td>5-year</td>
<td>0.16</td>
<td>1.35*</td>
<td>-0.08</td>
<td>0.86</td>
</tr>
<tr>
<td>7-year</td>
<td>-0.08</td>
<td>-2.37**</td>
<td>-0.04</td>
<td>0.26</td>
</tr>
<tr>
<td>10-year</td>
<td>-0.17</td>
<td>-5.51**</td>
<td>0.13</td>
<td>-1.04</td>
</tr>
<tr>
<td><strong>RMSE (basis points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>15.74</td>
<td>14.87</td>
<td>3.39</td>
<td>4.85</td>
</tr>
<tr>
<td>1-year</td>
<td>5.02</td>
<td>6.04</td>
<td>4.48</td>
<td>6.05</td>
</tr>
<tr>
<td>2-year</td>
<td>10.43</td>
<td>9.64</td>
<td>2.52</td>
<td>3.81</td>
</tr>
<tr>
<td>3-year</td>
<td>8.84</td>
<td>9.36</td>
<td>1.42</td>
<td>1.63</td>
</tr>
<tr>
<td>4-year</td>
<td>5.92</td>
<td>7.25</td>
<td>2.56</td>
<td>3.86</td>
</tr>
<tr>
<td>5-year</td>
<td>3.53</td>
<td>4.76</td>
<td>2.86</td>
<td>4.34</td>
</tr>
<tr>
<td>7-year</td>
<td>5.01</td>
<td>5.28</td>
<td>1.57</td>
<td>1.97</td>
</tr>
<tr>
<td>10-year</td>
<td>11.45</td>
<td>13.94</td>
<td>3.95</td>
<td>5.92</td>
</tr>
<tr>
<td><strong>autocorrelation of $\hat{e}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>0.98</td>
<td>0.97</td>
<td>0.92</td>
<td>0.95</td>
</tr>
<tr>
<td>1-year</td>
<td>0.75</td>
<td>0.88</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>2-year</td>
<td>0.90</td>
<td>0.90</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>3-year</td>
<td>0.88</td>
<td>0.91</td>
<td>0.86</td>
<td>0.89</td>
</tr>
<tr>
<td>4-year</td>
<td>0.79</td>
<td>0.89</td>
<td>0.91</td>
<td>0.95</td>
</tr>
<tr>
<td>5-year</td>
<td>0.52</td>
<td>0.83</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>7-year</td>
<td>0.82</td>
<td>0.88</td>
<td>0.82</td>
<td>0.88</td>
</tr>
<tr>
<td>10-year</td>
<td>0.95</td>
<td>0.97</td>
<td>0.91</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Figure 1
Actual and model-implied curvature
Table 5: Out-of-sample yield fits

<table>
<thead>
<tr>
<th></th>
<th>$A_1(2)$</th>
<th>$A_1(3)$ USV</th>
<th>$A_1(3)$</th>
<th>$A_1(4)$ USV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean $\hat{\epsilon}$ (basis points)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>5.25</td>
<td>-9.26**</td>
<td>-4.13**</td>
<td>-0.04</td>
</tr>
<tr>
<td>1-year</td>
<td>2.92**</td>
<td>-1.29**</td>
<td>4.73**</td>
<td>0.40</td>
</tr>
<tr>
<td>2-year</td>
<td>-2.52</td>
<td>4.87*</td>
<td>3.74**</td>
<td>-0.40</td>
</tr>
<tr>
<td>3-year</td>
<td>-5.30</td>
<td>5.84**</td>
<td>-0.93**</td>
<td>-0.58*</td>
</tr>
<tr>
<td>4-year</td>
<td>-5.39*</td>
<td>4.88**</td>
<td>-3.63**</td>
<td>-0.07</td>
</tr>
<tr>
<td>5-year</td>
<td>-3.72**</td>
<td>3.04**</td>
<td>-4.15**</td>
<td>0.62</td>
</tr>
<tr>
<td>7-year</td>
<td>1.78</td>
<td>-1.96*</td>
<td>-1.31**</td>
<td>1.23**</td>
</tr>
<tr>
<td>10-year</td>
<td>9.62*</td>
<td>-11.24**</td>
<td>5.41**</td>
<td>-1.16</td>
</tr>
<tr>
<td></td>
<td>RMSE (basis points)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>20.22</td>
<td>16.74</td>
<td>5.13</td>
<td>&gt;** 2.99</td>
</tr>
<tr>
<td>1-year</td>
<td>3.69</td>
<td>2.86</td>
<td>4.98</td>
<td>&gt;** 1.90</td>
</tr>
<tr>
<td>2-year</td>
<td>11.74</td>
<td>9.57</td>
<td>4.98</td>
<td>&gt;** 1.27</td>
</tr>
<tr>
<td>3-year</td>
<td>13.04</td>
<td>9.62</td>
<td>1.12</td>
<td>&gt;** 2.23</td>
</tr>
<tr>
<td>4-year</td>
<td>10.42</td>
<td>7.25</td>
<td>4.61</td>
<td>&gt;** 2.28</td>
</tr>
<tr>
<td>5-year</td>
<td>5.98</td>
<td>4.13</td>
<td>5.44</td>
<td>&gt;** 1.36</td>
</tr>
<tr>
<td>7-year</td>
<td>5.81</td>
<td>4.43</td>
<td>1.97</td>
<td>&gt;** 3.20</td>
</tr>
<tr>
<td>10-year</td>
<td>20.33</td>
<td>15.69</td>
<td>7.09</td>
<td>&gt;** 0.93</td>
</tr>
<tr>
<td></td>
<td>autocorrelation of $\hat{\epsilon}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>0.99</td>
<td>0.97</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>1-year</td>
<td>0.46</td>
<td>0.56</td>
<td>0.89</td>
<td>0.92</td>
</tr>
<tr>
<td>2-year</td>
<td>0.95</td>
<td>0.88</td>
<td>0.94</td>
<td>0.85</td>
</tr>
<tr>
<td>3-year</td>
<td>0.95</td>
<td>0.87</td>
<td>0.82</td>
<td>0.94</td>
</tr>
<tr>
<td>4-year</td>
<td>0.93</td>
<td>0.78</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>5-year</td>
<td>0.82</td>
<td>0.47</td>
<td>0.94</td>
<td>0.86</td>
</tr>
<tr>
<td>7-year</td>
<td>0.95</td>
<td>0.90</td>
<td>0.96</td>
<td>0.78</td>
</tr>
<tr>
<td>10-year</td>
<td>0.99</td>
<td>0.97</td>
<td>0.94</td>
<td>0.87</td>
</tr>
<tr>
<td>Correlation</td>
<td>$A_0(2)$</td>
<td>$A_1(3)$</td>
<td>$A_1(4)$</td>
<td>Daily</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>Actual vs. model average yield</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Actual vs. model slope</td>
<td>0.998</td>
<td>0.997</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>Actual vs. model curvature</td>
<td>0.254</td>
<td>0.382</td>
<td>0.998</td>
<td>0.997</td>
</tr>
<tr>
<td>Rolling vs. model volatility</td>
<td>0.771</td>
<td>-0.600</td>
<td>0.783</td>
<td>0.957</td>
</tr>
<tr>
<td>Eurodollar implied vs. model volatility</td>
<td>0.616</td>
<td>-0.498</td>
<td>0.605</td>
<td>0.693</td>
</tr>
<tr>
<td>Actual curvature vs. model volatility</td>
<td>-0.067</td>
<td>0.275</td>
<td>-0.087</td>
<td>-0.052</td>
</tr>
<tr>
<td>Actual curvature vs. model variance</td>
<td>-0.059</td>
<td>0.285</td>
<td>-0.072</td>
<td>-0.020</td>
</tr>
</tbody>
</table>
Figure 2
Rolling window and model-implied short rate volatility
Table 7: In-sample volatility forecasts

<table>
<thead>
<tr>
<th></th>
<th>$A_i(2)$</th>
<th>$A_i(3)$ USV</th>
<th>$A_i(3)$</th>
<th>$A_i(4)$ USV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bias in $\hat{\sigma}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>-6.20**</td>
<td>-2.82**</td>
<td>-4.16**</td>
<td>-2.02**</td>
</tr>
<tr>
<td>1-year</td>
<td>-3.27**</td>
<td>0.04</td>
<td>-2.77**</td>
<td>-1.02**</td>
</tr>
<tr>
<td>2-year</td>
<td>-1.11*</td>
<td>1.67**</td>
<td>-1.93**</td>
<td>-1.39**</td>
</tr>
<tr>
<td>3-year</td>
<td>-0.71</td>
<td>1.60**</td>
<td>-1.52**</td>
<td>-2.03**</td>
</tr>
<tr>
<td>4-year</td>
<td>-0.71</td>
<td>1.22*</td>
<td>-1.31**</td>
<td>-2.29**</td>
</tr>
<tr>
<td>5-year</td>
<td>-0.78</td>
<td>0.90</td>
<td>-1.20*</td>
<td>-2.27**</td>
</tr>
<tr>
<td>7-year</td>
<td>-0.90*</td>
<td>0.58</td>
<td>-1.10*</td>
<td>-1.89**</td>
</tr>
<tr>
<td>10-year</td>
<td>-0.97*</td>
<td>0.82*</td>
<td>-0.93*</td>
<td>-1.14**</td>
</tr>
<tr>
<td><strong>RMSE of $\hat{\sigma}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>8.23</td>
<td>&gt;**</td>
<td>5.78</td>
<td>6.95</td>
</tr>
<tr>
<td>1-year</td>
<td>6.91</td>
<td>&gt;**</td>
<td>5.89</td>
<td>&lt;*</td>
</tr>
<tr>
<td>2-year</td>
<td>6.91</td>
<td></td>
<td>7.01</td>
<td>7.19</td>
</tr>
<tr>
<td>3-year</td>
<td>6.81</td>
<td></td>
<td>7.06</td>
<td>7.03</td>
</tr>
<tr>
<td>4-year</td>
<td>6.63</td>
<td></td>
<td>6.85</td>
<td>6.79</td>
</tr>
<tr>
<td>5-year</td>
<td>6.47</td>
<td></td>
<td>6.64</td>
<td>6.60</td>
</tr>
<tr>
<td>7-year</td>
<td>6.24</td>
<td></td>
<td>6.31</td>
<td>6.34</td>
</tr>
<tr>
<td>10-year</td>
<td>6.08</td>
<td></td>
<td>6.07</td>
<td>6.17</td>
</tr>
<tr>
<td></td>
<td>$A_i(2)$</td>
<td>$A_i(3)$ USV</td>
<td>$A_i(3)$</td>
<td>$A_i(4)$ USV</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>--------------</td>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>bias in $\hat{\sigma}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>-9.60**</td>
<td>-2.97**</td>
<td>-8.22**</td>
<td>-3.69**</td>
</tr>
<tr>
<td>1-year</td>
<td>-6.18**</td>
<td>0.07</td>
<td>-6.37**</td>
<td>-2.41**</td>
</tr>
<tr>
<td>2-year</td>
<td>-2.63**</td>
<td>2.75**</td>
<td>-4.46**</td>
<td>-1.67**</td>
</tr>
<tr>
<td>3-year</td>
<td>-1.16</td>
<td>3.40**</td>
<td>-3.19**</td>
<td>-1.39*</td>
</tr>
<tr>
<td>4-year</td>
<td>-0.43</td>
<td>3.46**</td>
<td>-2.40**</td>
<td>-1.01</td>
</tr>
<tr>
<td>5-year</td>
<td>-0.12</td>
<td>3.28**</td>
<td>-2.01*</td>
<td>-0.69</td>
</tr>
<tr>
<td>7-year</td>
<td>-0.10</td>
<td>2.76**</td>
<td>-1.89*</td>
<td>-0.27</td>
</tr>
<tr>
<td>10-year</td>
<td>-0.50</td>
<td>2.34**</td>
<td>-2.13*</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>RMSE of $\hat{\sigma}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-month</td>
<td>9.97</td>
<td>&gt;**</td>
<td>&lt;**</td>
<td>8.66</td>
</tr>
<tr>
<td>1-year</td>
<td>7.22</td>
<td>&gt;**</td>
<td>&lt;**</td>
<td>7.40</td>
</tr>
<tr>
<td>2-year</td>
<td>5.93</td>
<td>&lt;**</td>
<td></td>
<td>7.01</td>
</tr>
<tr>
<td>3-year</td>
<td>6.18</td>
<td>&lt;*</td>
<td>7.15</td>
<td>6.97</td>
</tr>
<tr>
<td>4-year</td>
<td>6.43</td>
<td>&lt;*</td>
<td>7.44</td>
<td>7.00</td>
</tr>
<tr>
<td>5-year</td>
<td>6.56</td>
<td>&lt;*</td>
<td>7.45</td>
<td>7.02</td>
</tr>
<tr>
<td>7-year</td>
<td>6.49</td>
<td></td>
<td>7.13</td>
<td>6.92</td>
</tr>
<tr>
<td>10-year</td>
<td>6.11</td>
<td></td>
<td>6.56</td>
<td>6.60</td>
</tr>
</tbody>
</table>
Figure 3
The maturity/volatility relation
Volatility Forecast Regressions

5-day realized volatility = $\alpha + \beta \times$ forecast variable + $\epsilon$

Table 9A: Short rate volatility forecast regressions

<table>
<thead>
<tr>
<th>Specification Number</th>
<th>Intercept*</th>
<th>GARCH volatility</th>
<th>$A_t(3)$ forecast</th>
<th>$A_t(4)$ USV forecast</th>
<th>1st PC*</th>
<th>2nd PC*</th>
<th>3rd PC*</th>
<th>Adjusted R-Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>1'</td>
<td>-0.054</td>
<td></td>
<td></td>
<td></td>
<td>0.839</td>
<td>-0.437</td>
<td>1.228</td>
<td>0.155</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td></td>
<td></td>
<td></td>
<td>(0.504)</td>
<td>(0.076)</td>
<td>(0.541)</td>
<td></td>
</tr>
<tr>
<td>2'</td>
<td>0.000</td>
<td>0.697</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.077)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3'</td>
<td>-0.036</td>
<td>0.565</td>
<td></td>
<td></td>
<td>0.227</td>
<td>-0.175</td>
<td>0.770</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.084)</td>
<td></td>
<td></td>
<td>(0.326)</td>
<td>(0.065)</td>
<td>(0.309)</td>
<td></td>
</tr>
<tr>
<td>4'</td>
<td>0.007</td>
<td>-5.046</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.890)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'</td>
<td>7.695</td>
<td>-53.503</td>
<td></td>
<td></td>
<td>-14.607</td>
<td>2.986</td>
<td>-4.242</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(6.338)</td>
<td>(43.767)</td>
<td></td>
<td></td>
<td>(12.664)</td>
<td>(2.815)</td>
<td>(4.484)</td>
<td></td>
</tr>
<tr>
<td>6'</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td>0.698</td>
<td></td>
<td></td>
<td>0.238</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td>(0.073)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7'</td>
<td>-0.007</td>
<td>0.672</td>
<td></td>
<td></td>
<td>-0.525</td>
<td>-0.032</td>
<td>0.809</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.103)</td>
<td></td>
<td></td>
<td>(0.464)</td>
<td>(0.094)</td>
<td>(0.391)</td>
<td></td>
</tr>
</tbody>
</table>

* denotes a coefficient that has been multiplied by 100
Conclusions and summary

• Volatility in the standard three-factor model is incapable of fulfilling its dual role.
  – In our sample, estimates favor the cross-sectional role of volatility.
  – Fitted $A_1(3)$ volatility has no relation to actual.

• Since three factors in yields are needed, a fourth factor generating stochastic volatility is required.

• Yield curve evidence consistent with unspanned stochastic volatility.

• Consistent with other recent evidence of USV:
  – from implied option volatilities: Li Zhao (06), Duarte (06), Han (06), Jarrow, Li, Zhao (07), Trolle and Schwartz (07)
  – from realized yield volatility: Andersen Benzoni (06)
  – However, not unanimous: Fan, Gupta and Ritchken (03), Joslin (06).

• Complementary (but different?) to evidence (Cochrane and Piazzesi) of “unspanned risk premia.”