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 "$A_1(3)$" 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:
  - $Y_t = a_0 + a^\top X_t$
  - $\mathrm{Var}(dY_t) = b_0 + b^\top X_t$

  $\Rightarrow \mathrm{Var}(dY_t) = c_0 + c^\top Y_t$

- 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 "$A_1(3)$" 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:
  - $Y_t = a_0 + a^\top X_t$
  - $\mathrm{Var}(dY_t) = b_0 + b^\top X_t$

  $\Rightarrow \mathrm{Var}(dY_t) = c_0 + c^\top Y_t$

- 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.,

$$Y_t = a_0 + a^\top X_t \qquad BUT \qquad \mathrm{Var}(dY_t) \neq c_0 + c^\top Y_t$$

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:
  - $Y_t = a_0 + a^\top X_t$
  - $\mathrm{Var}(dY_t) = b_0 + b^\top X_t$

  $\Rightarrow \mathrm{Var}(dY_t) = c_0 + c^\top Y_t$

- 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.,

$$Y_t = a_0 + a^\top X_t \qquad BUT \qquad \mathrm{Var}(dY_t) \neq c_0 + c^\top Y_t$$

- 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)

$$\begin{aligned} dr &= \kappa_r(\theta_r - r)\,dt + \sqrt{V}\,dz_1^{\mathcal{Q}} \\ dV &= \kappa_V(\theta_V - V)\,dt + \sigma\sqrt{V}dz_2^{\mathcal{Q}} \end{aligned}$$

- In this model zero-coupon bond prices given by: $\boxed{P^T(t) = e^{A(T-t)+B(T-t)\,r_t+C(T-t)V_t}}$

$\Rightarrow$ Volatility risk <u>can</u> be hedged with appropriate position in any two bonds.

$\Rightarrow$ 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)

$$\begin{aligned} dr &= \kappa_r(\theta_r - r)\,dt + \sqrt{V}\,dz_1^{\mathcal{Q}} \\ dV &= \kappa_V(\theta_V - V)\,dt + \sigma\sqrt{V}dz_2^{\mathcal{Q}} \end{aligned}$$

- In this model zero-coupon bond prices given by: $\boxed{P^T(t) = e^{A(T-t)+B(T-t)\,r_t+C(T-t)V_t}}$

$\Rightarrow$ Volatility risk <u>can</u> be hedged with appropriate position in any two bonds.

$\Rightarrow$ 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):

$$\begin{aligned} \frac{dS}{S} &= r\,dt + \sqrt{V}\,dz_1^{\mathcal{Q}} \\ dV &= \kappa(\theta - V)\,dt + \sigma\sqrt{V}dz_2^{\mathcal{Q}} \end{aligned}$$

$\Rightarrow$ Volatility risk cannot be hedged by any portfolio of stock and bond

- Difference: equity models a **traded** asset, and **its** volatility.

What is 'Unspanned Stochastic Volatility'?

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

$$\begin{aligned} dr &= \kappa_r(\theta_r - r)\,dt + \sqrt{V}\,dz_1^{\mathcal{Q}} \\ dV &= \kappa_V(\theta_V - V)\,dt + \sigma\sqrt{V}dz_2^{\mathcal{Q}} \end{aligned}$$

- In this model zero-coupon bond prices given by: $\boxed{P^T(t) = e^{A(T-t)+B(T-t)\,r_t+C(T-t)V_t}}$

$\Rightarrow$ Volatility risk <u>can</u> be hedged with appropriate position in any two bonds.

$\Rightarrow$ 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):

$$\begin{aligned} \frac{dS}{S} &= r\,dt + \sqrt{V}\,dz_1^{\mathcal{Q}} \\ dV &= \kappa(\theta - V)\,dt + \sigma\sqrt{V}dz_2^{\mathcal{Q}} \end{aligned}$$

$\Rightarrow$ 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:

$$\frac{dP^T(t)}{P^T(t)} = r_t\,dt - \sigma_P(T-t)\sqrt{V_t}dz_1^Q(t)$$

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**

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

  where $\mathcal{K}^Q$ and Σ 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 Σ 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**

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

  where $\mathcal{K}^Q$ and Σ 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
  - μ^Q: the Q drift of r
  - θ^Q: the Q drift of μ^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
  - μ^Q: the Q drift of r
  - θ^Q: the Q drift of μ^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
  - μ^Q: the Q drift of r
  - θ^Q: the Q drift of μ^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:

$$\mathrm{E}\left[dX_t\right]/dt = a^Q + b^Q X_t$$

$$\mathrm{Cov}\left(dX_t, dX_t^\top\right)/dt = \Omega_0 + \sum_{i=1}^{m} \Omega_i X_{i,t}$$

where X is some combination of r, μ^Q, θ^Q, and/or V.

Example: the $A_1(3)$ model

If we define the state vector as

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

then

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

and

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

- 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:

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

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:

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

- Volatility affects long term yields via two channels:
  - Direct impact on the drift of μ^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:

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

- Volatility affects long term yields via two channels:
  - Direct impact on the drift of μ^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

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

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

$$\frac{1}{dt}\mathrm{Cov}\left(dX_t, dX_t^\top\right) = \begin{bmatrix} \underline{V} & \underline{c}_{r\mu} \\ \underline{c}_{r\mu} & \underline{\sigma}_\mu \end{bmatrix}$$

- Level and slope factors
- Constant covariance matrix
- 12 parameters total

Alternative: the $A_1(4)$ model

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

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

$$\frac{1}{dt}\mathrm{Cov}\left(dX_t, dX_t^\top\right) = \begin{bmatrix} \underline{V} & \underline{c}_{r\mu} & \underline{c}_{r\theta} & 0 \\ \underline{c}_{r\mu} & \underline{\sigma}_\mu & \underline{c}_{\mu\theta} & 0 \\ \underline{c}_{r\theta} & \underline{c}_{\mu\theta} & \underline{\sigma}_\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & c_{r\mu} & c_{r\theta} & c_{rV} \\ c_{r\mu} & \sigma_\mu & c_{\mu\theta} & c_{\mu V} \\ c_{r\theta} & c_{\mu\theta} & \sigma_\theta & c_{\theta V} \\ c_{rV} & c_{\mu V} & c_{\theta V} & \sigma_V \end{bmatrix} (V_t - \underline{V})$$

- 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.
- γ_V and κ_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:

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: $p(Y_t|X_{t-1})$

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

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:

$$dX_t \sim \mathrm{N}\left(\left(a + bX_t\right) dt,\ \left(\Omega_0 + \Omega_V (V_t - \underline{V})\right) dt \right)$$

discretized using the Euler approximation.

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

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:

$$dX_t \sim \mathrm{N}\left(\left(a + bX_t\right) dt,\ \left(\Omega_0 + \Omega_V (V_t - \underline{V})\right) dt \right)$$

discretized using the Euler approximation.

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

- 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), Lamoureaux and Witte (2002), Bester (2004), Sanford and Martin (2003), and Polson, Stroud, and Müller (2001)

- Diffusion process:

$$dX_t \sim \mathrm{N}\left(\left(a + bX_t\right) dt,\ \left(\Omega_0 + \Omega_V(V_t - \underline{V})\right) dt\right)$$

  discretized using the Euler approximation.

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

- 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.

- 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:

$$dX_t \sim \mathrm{N}\left(\left(a + bX_t\right) dt,\ \left(\Omega_0 + \Omega_V(V_t - \underline{V})\right) dt\right)$$

  discretized using the Euler approximation.

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

- 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.

- 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

- Libor and swap rates from 1988 to 2005.
- 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 2A: Posterior distributions of model parameters

	$A_0(2)$	$A_1(3)$ USV	$A_1(3)$	$A_1(4)$ USV
m_0	0.003	0.003	0.211	
	(0.002, 0.003)	(0.002, 0.003)	(0.188, 0.238)	
m_r	-0.019	*-0.031*	-1.238	
	(-0.025, -0.014)	*(-0.032, -0.028)*	(-1.348, -1.131)	
m_μ	-0.633	*-0.371*	-2.804	
	(-0.705, -0.585)	*(-0.382, -0.354)*	(-2.959, -2.597)	
m_V		*1.000*	-923.149	
			(-1003.107, -811.230)	
$\gamma_V^P \times 10^4$		1.015	0.568	0.803
		(0.601, 1.899)	(0.126, 1.018)	(0.469, 1.147)
κ_V^P		1.130	0.441	1.014
		(2.537, 0.484)	(0.756, 0.056)	(2.116, 0.412)
$\underline{V} \times 10^4$	1.021	0.020	0.111	0.003
	(0.883, 1.193)	(0.003, 0.041)	(0.046, 0.298)	(0.000, 0.011)
$\underline{\sigma}_\mu \times 10^4$	0.981	0.182	5.809	3.448
	(0.643, 1.497)	(0.158, 0.207)	(3.208, 11.349)	(2.675, 3.999)
$\underline{c}_{r\mu} \times 10^4$	-0.376	0.043	0.708	-0.102
	(-0.637, -0.240)	(0.017, 0.071)	(0.187, 0.926)	(-0.197, -0.014)
σ_μ		*0.015*	17.660	*0.008*
		(0.014, 0.016)	(11.327, 24.871)	*(0.007, 0.008)*
$\sigma_V \times 10^4$		1.647	0.025	1.221
		(0.948, 2.845)	(0.022, 0.030)	(0.811, 1.968)
$c_{r\mu}$		-0.124	-3.617	-0.087
		(-0.127, -0.118)	(-4.526, -2.893)	(-0.091, -0.083)
$c_{rV} \times 10^3$		-0.003	-0.376	2.685
		(-2.430, 3.002)	(-0.504, -0.209)	(0.152, 4.457)
$c_{\mu V}$		*0.000*	-1.327	*-0.233*

Specification analysis

In the Euler approximation

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

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

Table 3: Specification analysis

	$A_0(2)$		$A_1(3)$ USV			$A_1(3)$			$A_1(4)$ USV			
	r	μ^Q	r	μ^Q	V	r	μ^Q	V	r	μ^Q	θ^Q	V
Mean $\hat{\epsilon}$	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	-0.01
SD of $\hat{\epsilon}$	0.97	0.99	1.00	1.12	0.99*	0.95	0.96*	0.98	0.99	0.98	0.98	0.99
Skewness of $\hat{\epsilon}$	-0.80*	0.31	-0.03	0.04	0.06	-0.35	0.03*	-0.06	-0.02	-0.02	0.01	0.07
Kurtosis of $\hat{\epsilon}$	7.42*	3.93	3.62	4.13*	2.96*	7.21*	4.04*	4.19*	4.32	4.06*	4.08*	2.97*
1st order AC of $\hat{\epsilon}$	0.19*	0.24	0.10	-0.02*	-0.01	0.05*	0.04*	-0.12*	0.05	0.03*	0.03*	0.00*
1st order AC of $\lvert\hat{\epsilon}\rvert$	0.14*	0.07	0.00	-0.01	0.00	0.22*	0.10*	0.06*	-0.01	0.10	0.09*	0.00*
1st order AC of $\hat{\epsilon}^2$	0.05*	0.04	-0.01	-0.01	0.00	0.16*	0.08*	0.06*	-0.01	0.08	0.08*	0.00*

Table 4: In-sample yield fits

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

Figure 1

Actual and model-implied curvature

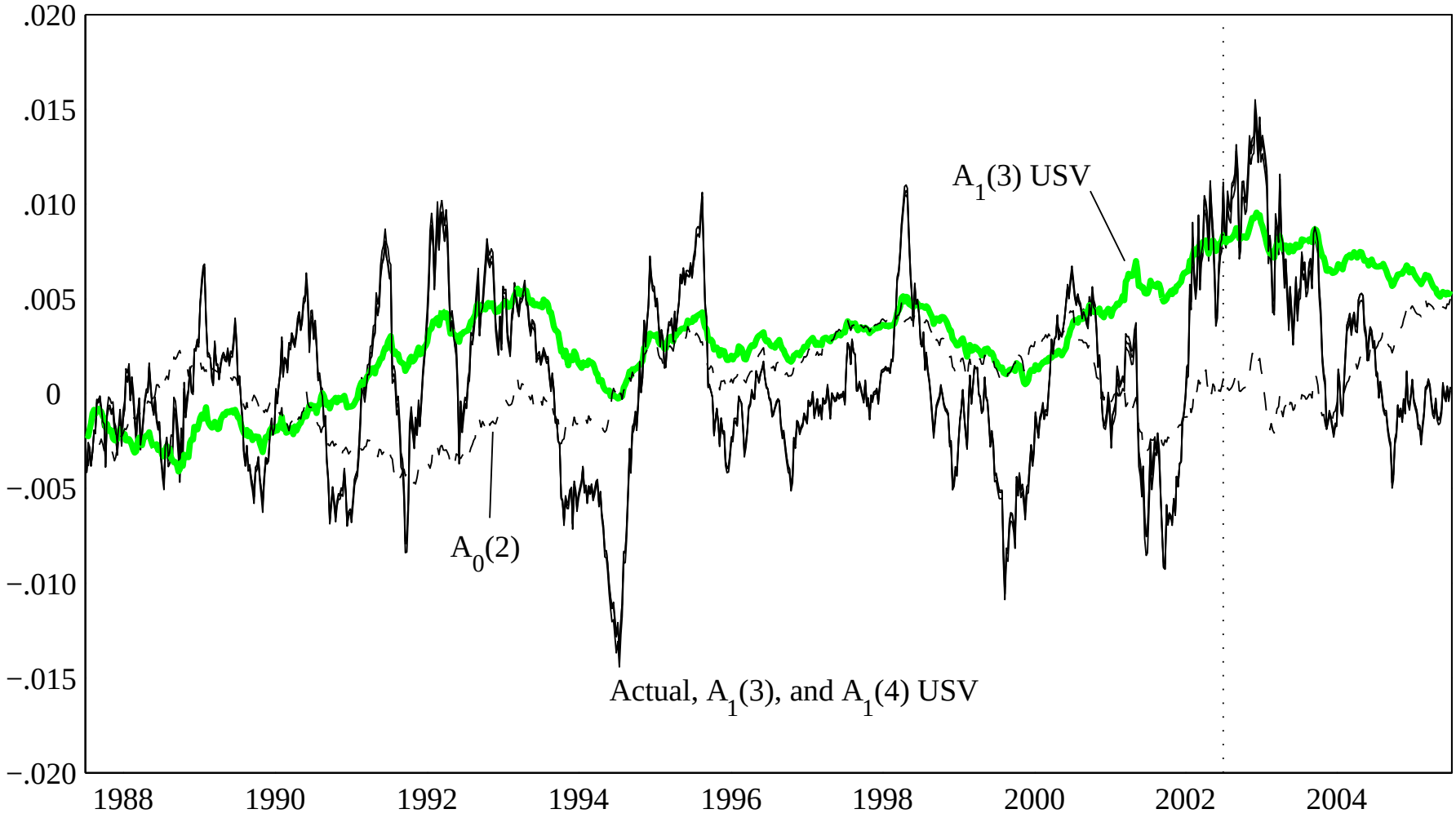


Table 5: Out-of-sample yield fits

	$A_0(2)$	$A_1(3)$ USV		$A_1(3)$		$A_1(4)$ USV
			mean $\hat{e}$ (basis points)			
6-month	5.25	-9.26**		-4.13**		-0.04
1-year	2.92**	-1.29**		4.73**		0.40
2-year	-2.52	4.87*		3.74**		-0.40
3-year	-5.30	5.84**		-0.93**		-0.58*
4-year	-5.39*	4.88**		-3.63**		-0.07
5-year	-3.72**	3.04**		-4.15**		0.62
7-year	1.78	-1.96*		-1.31**		1.23**
10-year	9.62*	-11.24**		5.41**		-1.16
			RMSE (basis points)			
6-month	20.22	16.74	>**	5.13	>**	2.99
1-year	3.69	2.86	<**	5.77	>*	3.98
2-year	11.74	9.57	>**	4.98	>**	1.90
3-year	13.04	9.62	>**	1.12		1.27
4-year	10.42	7.25	>*	4.61	>**	2.23
5-year	5.98	4.13		5.44	>**	2.28
7-year	5.81	4.43	>**	1.97	>*	1.36
10-year	20.33	15.69	>**	7.09	>**	3.20
			autocorrelation of $\hat{e}$			
6-month	0.99	0.97		0.91		0.91
1-year	0.46	0.56		0.89		0.92
2-year	0.95	0.88		0.94		0.85
3-year	0.95	0.87		0.82		0.94
4-year	0.93	0.78		0.93		0.89
5-year	0.82	0.47		0.94		0.86
7-year	0.95	0.90		0.96		0.78
10-year	0.99	0.97		0.94		0.87

Table 6: Correlations of observed and model-implied time series

	$A_0(2)$	$A_1(3)$ USV	$A_1(3)$	$A_1(4)$ USV	Daily GARCH	Implied vol Euro$	Implied vol Cap/Floor
Actual vs. model average yield	1.000	1.000	1.000	1.000			
Actual vs. model slope	0.998	0.997	0.998	0.998			
Actual vs. model curvature	0.254	0.382	0.998	0.997			
Rolling vs. model volatility		0.771	-0.600	0.783	0.957	0.676	0.548
Eurodollar implied vs. model volatility		0.616	-0.498	0.605	0.693	1.000	0.684
Actual curvature vs. model volatility		-0.067	0.275	-0.087	-0.052	-0.103	0.105
Actual curvature vs. model variance		-0.059	0.285	-0.072	-0.020	-0.124	0.102

Figure 2

Rolling window and model-implied short rate volatility

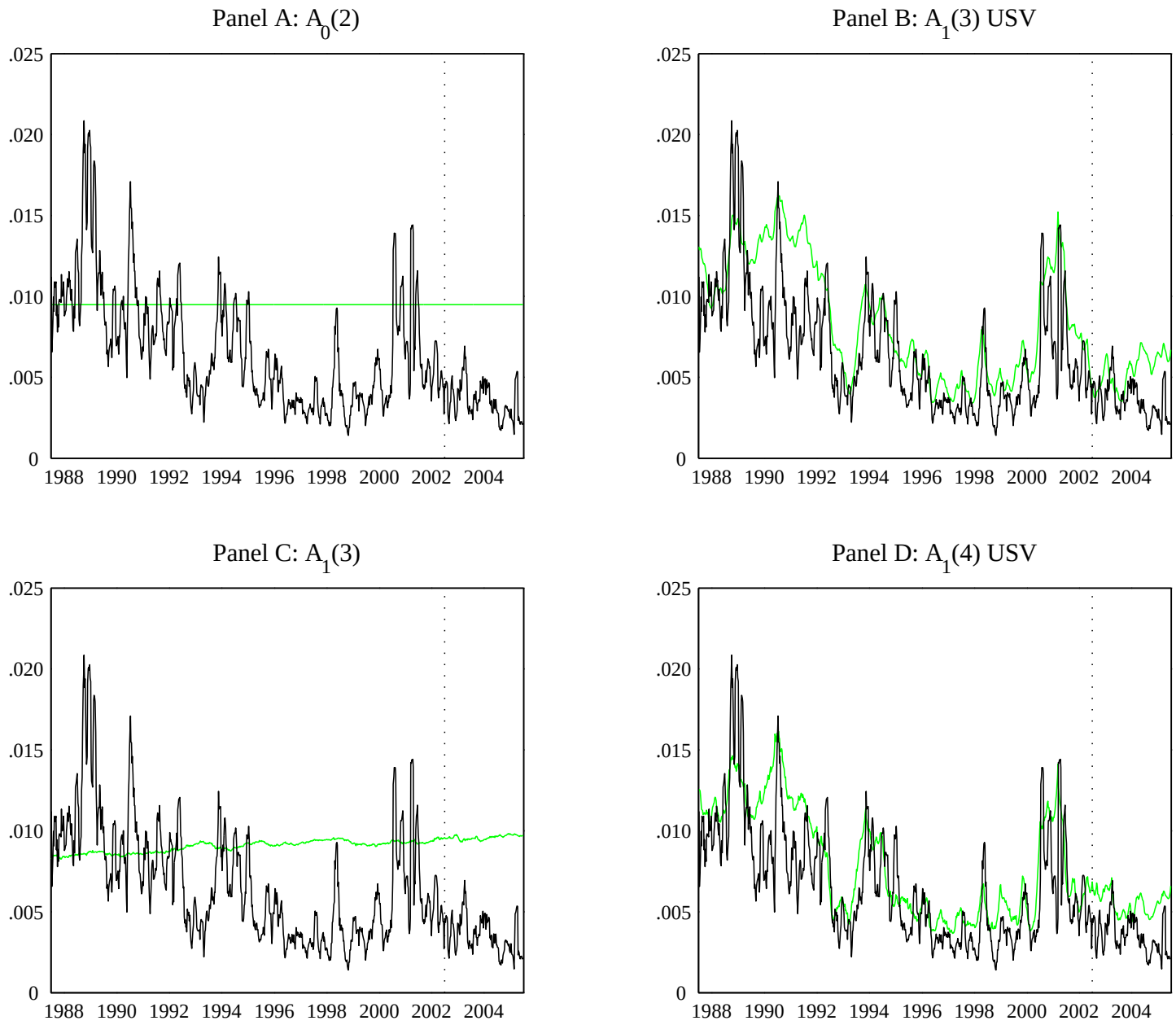


Table 7: In-sample volatility forecasts

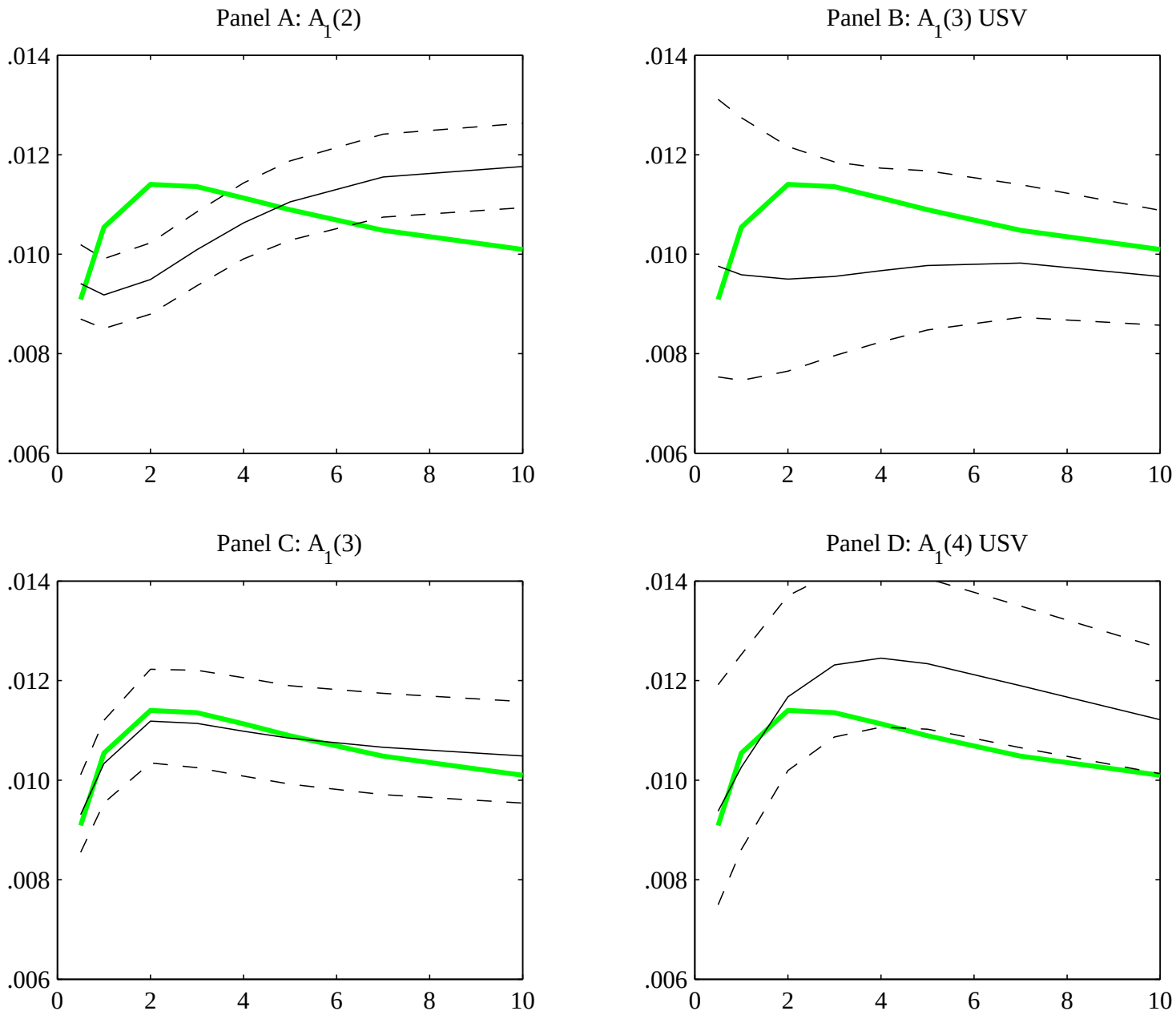
	$A_0(2)$		$A_1(3)$ USV		$A_1(3)$		$A_1(4)$ USV
				bias in $\hat{\sigma}$			
6-month	-6.20**		-2.82**		-4.16**		-2.02**
1-year	-3.27**		-0.04		-2.77**		-1.02**
2-year	-1.11*		1.67**		-1.93**		-1.39**
3-year	-0.71		1.60**		-1.52**		-2.03**
4-year	-0.71		1.22*		-1.31**		-2.29**
5-year	-0.78		0.90		-1.20*		-2.27**
7-year	-0.90*		0.58		-1.10*		-1.89**
10-year	-0.97*		0.82*		-0.93*		-1.14**
				RMSE of $\hat{\sigma}$			
6-month	8.23	>**	5.78	<*	6.95	>**	5.26
1-year	6.91	>**	5.89	<*	6.79	>**	5.87
2-year	6.91		7.01		7.19		6.87
3-year	6.81		7.06		7.03		7.09
4-year	6.63		6.85		6.79		7.05
5-year	6.47		6.64		6.60		6.91
7-year	6.24		6.31		6.34		6.54
10-year	6.08		6.07		6.17		6.12

Table 8: Out-of-sample volatility forecasts

	$A_0(2)$		$A_1(3)$ USV		$A_1(3)$		$A_1(4)$ USV
				bias in $\hat{\sigma}$			
6-month	-9.60**		-2.97**		-8.22**		-3.69**
1-year	-6.18**		0.07		-6.37**		-2.41**
2-year	-2.63**		2.75**		-4.46**		-1.67**
3-year	-1.16		3.40**		-3.19**		-1.39*
4-year	-0.43		3.46**		-2.40**		-1.01
5-year	-0.12		3.28**		-2.01*		-0.69
7-year	-0.10		2.76**		-1.89*		-0.27
10-year	-0.50		2.34**		-2.13*		0.03
				RMSE of $\hat{\sigma}$			
6-month	9.97	>**	4.32	<**	8.66	>**	4.69
1-year	7.22	>**	4.13	<**	7.40	>**	4.54
2-year	5.93		6.24		7.01	>**	5.64
3-year	6.18	<*	7.15		6.97	>**	6.27
4-year	6.43	<*	7.44		7.00	>**	6.53
5-year	6.56	<*	7.45		7.02	>*	6.62
7-year	6.49		7.13		6.92	>*	6.50
10-year	6.11		6.56		6.60	>*	6.08

Figure 3

The maturity/volatility relation



Volatility Forecast Regressions

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

Table 9A: Short rate volatility forecast regressions

Specification Number	Intercept*	GARCH volatility	$A_1(3)$ forecast	$A_1(4)$ USV forecast	1st PC*	2nd PC*	3rd PC*	Adjusted R-Squared
				dependent variable = $\hat{\sigma}$				
1'	-0.054				0.839	-0.437	1.228	0.155
	(0.022)				(0.504)	(0.076)	(0.541)	
2'	0.000	0.697						0.246
	(0.000)	(0.077)						
3'	-0.036	0.565			0.227	-0.175	0.770	0.268
	(0.012)	(0.084)			(0.326)	(0.065)	(0.309)	
4'	0.007		-5.046					0.135
	(0.001)		(0.890)					
5'	7.695		-53.503		-14.607	2.986	-4.242	0.157
	(6.338)		(43.767)		(12.664)	(2.815)	(4.484)	
6'	0.000			0.698				0.238
	(0.000)			(0.073)				
7'	-0.007			0.672	-0.525	-0.032	0.809	0.245
	(0.016)			(0.103)	(0.464)	(0.094)	(0.391)	

* 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."