Discussion of: "Estimating Exponential Affine Models with Correlated Measurement Errors" by M.A.H Dempster and Ke Tang

Jakub W. Jurek

Princeton University – Bendheim Center for Finance

September 2010

Paper Summary

Question: What is the effect of incorporating "realistic" error structures on estimates of affine term-structure models?

- Existing literature generally assumes i.i.d. measurement errors (Duan and Simonato (1999), Chen and Scott (2003), Schwartz (1997)).
- *Applications:* Bonds and commodities.

Method: Maximum likelihood estimation + (extended) Kalman filter.

- Requires specifying dynamics of latent factors (state equation) and structure of measurement errors (observation equation).

Paper Summary

Question: What is the effect of incorporating "realistic" error structures on estimates of affine term-structure models?

- Existing literature generally assumes i.i.d. measurement errors (Duan and Simonato (1999), Chen and Scott (2003), Schwartz (1997)).
- *Applications:* Bonds and commodities.

Method: Maximum likelihood estimation + (extended) Kalman filter.

- Requires specifying dynamics of latent factors (state equation) and structure of measurement errors (observation equation).

Results:

1. Incorporating error structure is important for inference regarding the dynamics of latent factors in affine term-structure models (ATSM).
2. Monte Carlo analysis indicates cross-sectional correlation is less of a problem, than time-series autocorrelation.
3. Likelihood ratio tests indicate the augmented state space form (SSF) is significantly better than the basic SSF.

Residuals from the fitting standard term-structure models are auto-correlated both in the cross-section and the time-series (de Jong (2000); this paper, Table 4).

	Covariance matrix									DW	Serial correlation
Basic SSF											
3 months	1.00	0.97	0.69	-0.18	-0.52	0.27	0.37	0.15	-0.43	**0.11**	**0.94** (0.010)
6 months	0.97	1.00	0.82	-0.04	-0.60	0.19	0.36	0.24	-0.45	**0.19**	**0.91** (0.013)
12 months	0.69	0.82	1.00	0.34	-0.66	-0.03	0.22	0.41	-0.39	**0.64**	**0.68** (0.022)
24 months	-0.18	-0.04	0.34	1.00	-0.31	-0.52	-0.27	0.52	0.02	**0.53**	**0.73** (0.021)
36 months	-0.52	-0.60	-0.66	-0.31	1.00	-0.48	-0.54	-0.17	0.54	**0.70**	**0.65** (0.023)
48 months	0.27	0.19	-0.03	-0.52	-0.48	1.00	0.46	-0.50	-0.23	**1.00**	**0.50** (0.026)
60 months	0.37	0.36	0.22	-0.27	-0.54	0.46	1.00	-0.42	-0.27	**0.51**	**0.73** (0.020)
72 months	0.15	0.24	0.41	0.52	-0.17	-0.50	-0.42	1.00	-0.59	**0.60**	**0.70** (0.022)
120 months	-0.43	-0.45	-0.39	0.02	0.54	-0.23	-0.27	-0.59	1.00	**0.33**	**0.84** (0.017)

- Breusch-Godfrey serial correlation LM test is a more robust alternative to the Durbin-Watson statistic.

Affine Term-Structure Models (ATSM)

In the canonical representation of an affine term structure model, the interest rate is driven by an N-factor state vector, Y_t:

$$\begin{aligned} r_t &= \phi_0 + \phi_Y' \cdot Y_t \\ dY_t &= K \cdot (\Theta - Y_t) \cdot dt + \Sigma \cdot \sqrt{\Upsilon_t} \cdot dW_t^{\mathbb{Q}} \end{aligned}$$

with variance terms, $[\Upsilon_t]_{ij} = a_{ij} + b_{ij} \cdot Y_t$.

Affine Term-Structure Models (ATSM)

In the canonical representation of an affine term structure model, the interest rate is driven by an N-factor state vector, Y_t:

$$\begin{aligned} r_t &= \phi_0 + \phi_Y' \cdot Y_t \\ dY_t &= K \cdot (\Theta - Y_t) \cdot dt + \Sigma \cdot \sqrt{\Upsilon_t} \cdot dW_t^{\mathbb{Q}} \end{aligned}$$

with variance terms, $[\Upsilon_t]_{ij} = a_{ij} + b_{ij} \cdot Y_t$.

- Bond prices are exponential affine in the state vector $\rightarrow$ yields are affine in Y_t.

$$\begin{aligned} y_t &= A(\tau) + B(\tau) \cdot Y_t \\ \Delta y_t &= B(\tau) \cdot \Delta Y_t \end{aligned}$$

Affine Term-Structure Models (ATSM)

In the canonical representation of an affine term structure model, the interest rate is driven by an N-factor state vector, Y_t:

$$\begin{aligned} r_t &= \phi_0 + \phi_Y' \cdot Y_t \\ dY_t &= K \cdot (\Theta - Y_t) \cdot dt + \Sigma \cdot \sqrt{\Upsilon_t} \cdot dW_t^{\mathbb{Q}} \end{aligned}$$

with variance terms, $[\Upsilon_t]_{ij} = a_{ij} + b_{ij} \cdot Y_t$.

- Bond prices are exponential affine in the state vector $\rightarrow$ yields are affine in Y_t.

$$\begin{aligned} y_t &= A(\tau) + B(\tau) \cdot Y_t \\ \Delta y_t &= B(\tau) \cdot \Delta Y_t \end{aligned}$$

- Dai and Singleton (2000) characterize the maximally flexible models within each class (canonical representation), $A_m(N)$ $(0 \leq m \leq N)$.
  - Examples: Vasicek $\in A_0(1)$, Cox-Ingersoll-Ross (1985) $\in A_1(1)$.

Principal Components

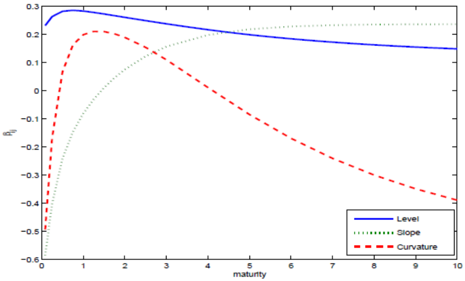
Litterman and Scheinkman (1991) argue that $\approx 99\%$ of the variation in yields can be explained using three factors.

- Principal components analysis
  - k linearly-independent series $\longrightarrow$ k factors.
  - Factors are orthogonal and maximize explanatory R^2.

Principal Components

Litterman and Scheinkman (1991) argue that ≈ 99% of the variation in yields can be explained using three factors.

- Principal components analysis
  - k linearly-independent series → k factors.
  - Factors are orthogonal and maximize explanatory R^2.
- Each factor can be represented as a linear combination of the underlying series (yields).



- U.S. data most consistent with $A_1(3)$ and $A_2(3)$ term structures (Dai and Singleton (2002)).

Principal Components

On the importance of squiggles

- **Data:** Jan 1982 - June 2010; Federal Reserve H15 Report (3M, 6M, 1Y, 2Y, 3Y, 5Y, 7Y, 10Y).

- Fraction of variance explained by principal components

	Levels ($y_{i,t}$)	Changes ($\Delta y_{i,t}$)
Factor 1 (level)	97.63%	78.52%
Factor 2 (slope)	2.23%	14.49%
Factor 3 (curvature)	0.11%	3.04%
Factor 4	0.02%	1.83%
Factors 5-8	$<$ 0.01%	2.11%

- Elementary my dear Watson ... the fourth factor is the *squiggle*.

	3m	6m	1y	2y	3y	5y	7y	10y
Levels	-0.46	0.39	0.51	-0.20	-0.32	-0.22	-0.10	0.42
Changes	-0.21	0.56	-0.07	-0.51	-0.37	0.00	0.27	0.39

Suppose the factors, Y_t, from the affine term-structure model, map onto the principal components. What should we expect to see from the residuals from a three-factor model?

- The residuals from projecting a k yields onto N-factors – an $A_m(N)$ model – would be linear combinations of the remaining $k - N$ principal components.

Principal Components

On the importance of squiggles

Suppose the factors, Y_t, from the affine term-structure model, map onto the principal components. What should we expect to see from the residuals from a three-factor model?

- The residuals from projecting a k yields onto N-factors – an $A_m(N)$ model – would be linear combinations of the remaining $k - N$ principal components.
- Compare the properties of the "model-fitting errors" with the *squiggle* factor.
  1. The fourth *level* factor is quite persistent, $\rho_1 = 0.95$ (daily) $\longrightarrow$ half-life of approximately one-month.
  2. Even the 8^{th} factor's persistence is $\rho_1 = 0.85$ (daily) $\longrightarrow$ half-life of about one week.

Suppose the factors, Y_t, from the affine term-structure model, map onto the principal components. What should we expect to see from the residuals from a three-factor model?

- The residuals from projecting a k yields onto N-factors – an $A_m(N)$ model – would be linear combinations of the remaining $k - N$ principal components.
- Compare the properties of the "model-fitting errors" with the *squiggle* factor.
  1. The fourth *level* factor is quite persistent, $\rho_1 = 0.95$ (daily) $\longrightarrow$ half-life of approximately one-month.
  2. Even the 8^{th} factor's persistence is $\rho_1 = 0.85$ (daily) $\longrightarrow$ half-life of about one week.
- *Prediction:* The errors from the augmented estimation procedure should still exhibit autocorrelation.

Suppose the factors, Y_t, from the affine term-structure model, map onto the principal components. What should we expect to see from the residuals from a three-factor model?

- The residuals from projecting a k yields onto N-factors – an $A_m(N)$ model – would be linear combinations of the remaining $k - N$ principal components.
- Compare the properties of the "model-fitting errors" with the *squiggle* factor.
  1. The fourth *level* factor is quite persistent, $\rho_1 = 0.95$ (daily) $\longrightarrow$ half-life of approximately one-month.
  2. Even the 8^{th} factor's persistence is $\rho_1 = 0.85$ (daily) $\longrightarrow$ half-life of about one week.
- *Prediction:* The errors from the augmented estimation procedure should still exhibit autocorrelation.

But ... what exactly is the empirical relation between the principal components and the factors?

- Do regressions of yields on the filtered factor series, Y_t, produce R^2 of 99%?

Errors vs. Factors

- **Recall:** Fit $A_m(3)$ model $\rightarrow$ examine fitting errors $\rightarrow$ add a factor structure to observation equation.

- How is incorporating a realistic structure for a model-fitting error different from adding another factor, $A_m(N+1)$?
  - Use likelihood ratio test to examine the whether the $A_m(N+1)$ improves on $A_m(N)$.
  - What is the influence of adding a fourth factor on the coefficient estimates of the first three factors?
  - What is the relation between the first principal component of the error series from estimating a three-factor model via ASSF, and the *fourth* factor in an $A_1(4)$ or $A_2(4)$?

Other Comments

1. What is the influence of using the discretized transition density, rather than the true density on parameter estimation?
   - Transition density not only depends on Y_t, but will generally be non-Gaussian.
   - Solves problem of positive-definiteness.

2. The effect of measurement errors on Kalman filter parameter estimates, likely depends on the structure of the noise vis-a-vis the structure of "actual" latent factors.
   - How robust is the conclusion that cross-sectional correlation is less important than time-series auto-correlation?
   - The error structure in the Monte Carlo analysis looks like a *convexity* factor.