Jointly Modelling and Robust Forecasting High-Dimensional Yield Curves

Wolfgang K. Härdle
Chen Huang
Linlin Niu

Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin
Wang Yanan Institute for Studies in Economics
Xiamen University
http://lvb.wiwi.hu-berlin.de
http://irtg1792.hu-berlin.de
http://wise.xmu.edu.cn/english

Yield Curves Data

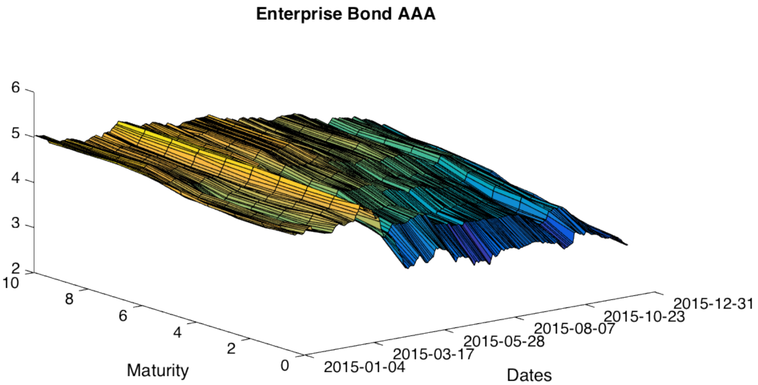


Figure 1: Daily yield curves of Chinese enterprise bond AAA in 2015.

Yield Curves Modelling

- Based on economic theory
  - market equilibrium: Vasicek model; CIR model
  - no-arbitrage: derivative pricing under B-S framework
  - affine-class: dynamic in maturities with time series technique
- Goodness of fit and forecasting
  - dynamic Nelson-Siegel model (Diebold and Li, 2006)
  - other generalized N-S models

Dynamic Nelson-Siegel Model

Advantages

- excellent fit to the term structure
- clear explanation on factors: level, slope and curvature
- estimation simplicity

Limitations

- specification issues
- jointly modelling across bond types and credit ratings

Go beyond DNS

- high-dimensional curves across types and ratings
- flexible representation through high-dimensional B-splines
- sparse latent factors
- robust estimation via LAD regression
- risky bonds with low credit ratings

Estimation Issues

- estimate a high-dimensional coefficient matrix
- nuclear norm penalty
  - involve a convex optimization
  - lead to a low dimensional factor model
- SVD to identify factors and loadings
- multivariate factorisable quantile/expectile regression (Chao et al. 2015; 2016)

Objectives and Contributions

- jointly modelling and robust forecasting high-dimensional yield curves
- multivariate factorisable median regression (MFMR)
- application for Chinese bond market
  - systemic liquidity and dispersion measure among curves
  - term structure and credit risk structure
  - in- and out-of-sample performance

Outline

1. Motivation ✓
2. Model and Estimation
3. Application with Chinese Yield Curve Data
4. Concluding Remarks

Model Specification

- $\mathbf{Y} = (Y_{ij}) \in \mathbb{R}^{n \times m}$: multivariate curves
  - m: the number of curves (across credit ratings and types)
  - n: the length of observations (over time)
- $\{\boldsymbol{X}_i\}_{i=1}^{n} \in \mathbb{R}^p$: B-spline basis functions
- $\max\{p, m\} \ll n$ while $p, m \to \infty$ is allowed
- refer to Chao et al. (2016) for more assumptions

Model Specification

- Linear sparse factor structure:

$$Y_{ij} = \sum_{k=1}^{r} \psi_{j,k} f_k(\boldsymbol{X}_i) + u_{ij}, \tag{1}$$

  where $f_k(\boldsymbol{X}_i)$ is the kth factor, r is the number of factors, $\psi_{j,k}$ are the factor loadings.

- Factors are constructed by linear combination of $\boldsymbol{X}_i$:

$$f_k(\boldsymbol{X}_i) = \boldsymbol{\varphi}_{\boldsymbol{k}}^{\top} \boldsymbol{X}_i \tag{2}$$

Model Specification

- Substituting (2) into (1):

$$Y_{ij} = \boldsymbol{\gamma}_j^\top \boldsymbol{X}_i + u_{ij}, \tag{3}$$

where $\boldsymbol{\gamma}_j = (\sum_{k=1}^{r} \psi_{j,k}\varphi_{k,1}, \ldots, \sum_{k=1}^{r} \psi_{j,k}\varphi_{k,p})^\top$

- To estimate the coefficient matrix $\boldsymbol{\Gamma}$, where $\boldsymbol{\gamma}_j$ is the j-th column of $\boldsymbol{\Gamma}$

Estimation

- Robust estimation on $\boldsymbol{\Gamma}$ via median regression

$$\widehat{\boldsymbol{\Gamma}} = \arg \min_{\boldsymbol{\Gamma} \in \mathbb{R}^{p \times m}} \left\{ (mn)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m} \left| Y_{ij} - \boldsymbol{X}_i^{\top} \boldsymbol{\Gamma}_{\cdot j} \right| + \lambda \|\boldsymbol{\Gamma}\|_* \right\} \tag{4}$$

  - nuclear norm $\|\boldsymbol{\Gamma}\|_* = \sum_{j=1}^{\min(p,m)} \sigma_j(\boldsymbol{\Gamma})$, given the singular values of $\boldsymbol{\Gamma}$: $\sigma_1(\boldsymbol{\Gamma}) \geq \sigma_2(\boldsymbol{\Gamma}) \geq \ldots \geq \sigma_{\min(p,m)}(\boldsymbol{\Gamma})$,
  - # of nonzero singular values of $\boldsymbol{\Gamma}$ is # of factors: r
  - smooth fast iterative shrinkage thresholding algorithm
  - singular value decomposition on $\boldsymbol{\Gamma}$

Data

- daily yield spread in Chinese bond market
- 180 spread curves
  - maturities of $1, 2, \ldots, 10$ years
  - enterprise bonds (9 credit ratings), chengtou bonds (5 credit ratings), company bonds (4 credit ratings)
- 733 observations from 2014.01 to 2016.12
- obtained from Wind Datafeed Service (WDS)

Factor Analysis

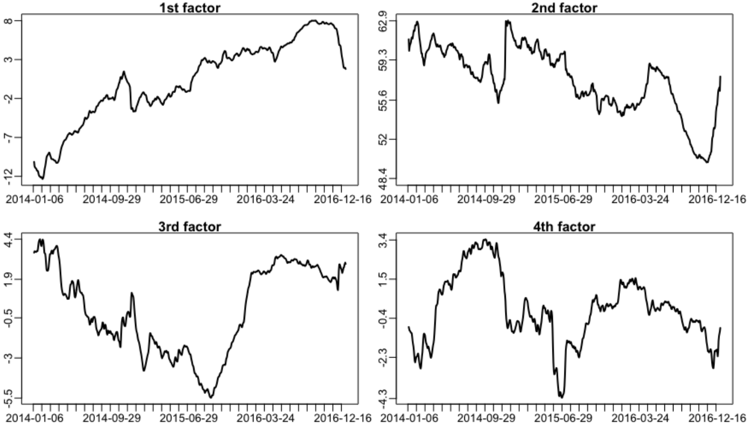


Figure 2: Plot of the first four factors (92.27% of the variance is explained).

Factor Loadings

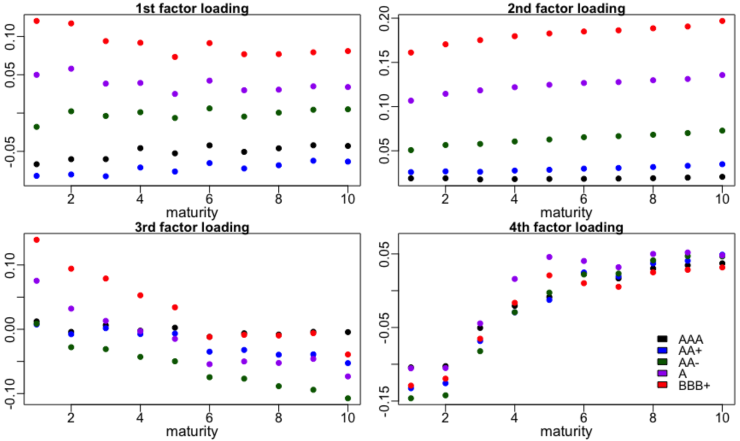


Figure 3: Factor loadings for enterprise bonds of five credit ratings.

Three Factors by DNS (Treasury Bond)

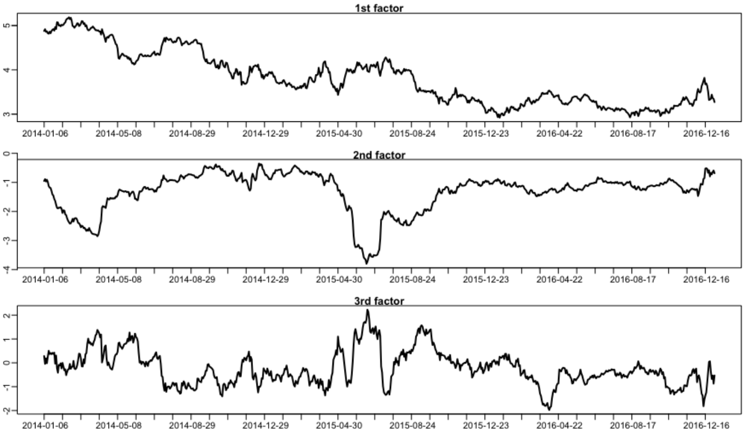


Figure 4: Plot of the three factors by DNS for treasury bond.

Factor Analysis

- Factors interpretation:
  - 1st: systemic liquidity or dispersion measure among curves - 53.49%
  - 2nd: level (credit risk) - 18.95%
  - 3rd: slope - 14.42%
  - 4th: curvature - 5.41%

Alternative Approaches

- Three factors DNS

$$Y_{i\tau} = f_{1i} + \left\{ \frac{1-\exp(-\lambda\tau)}{\lambda\tau} \right\} f_{2i} + \left\{ \frac{1-\exp(-\lambda\tau)}{\lambda\tau} - \exp(-\lambda\tau) \right\} f_{3i} + u_{i\tau},$$

  where τ denotes the maturities (for a particular bond type and credit rating).

- PCA

$$Y_{ij} = \sum_{k=1}^{r} \psi_{kj} f_{ki} + u_{ij},$$

  where $f_{k\cdot}^{\top} = \mathbf{Y}\gamma_k$, γ_k is the k-th eigenvector of $\mathrm{Var}(\mathbf{Y})$. VAR is applied to model the dynamics in factors.

Fitting Performance

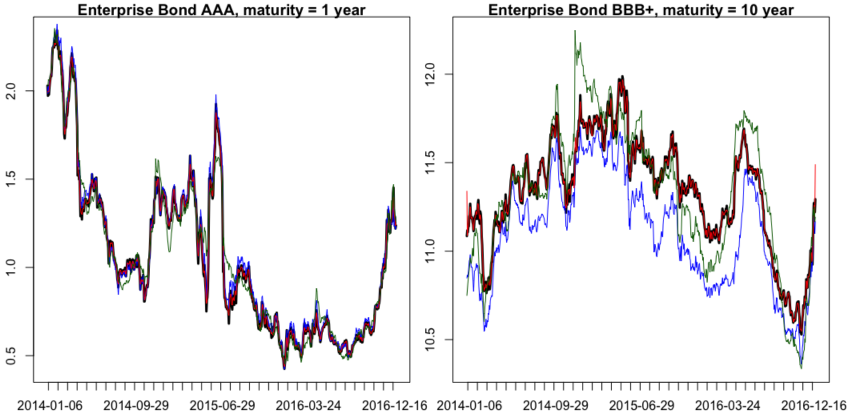


Figure 5: Fitted curves by MFMR, DNS, PCA, with real observations.

Fitting Performance - Whole Sample

		MFMR	DNS	PCA
Enterprise Bonds	AAA	1.92	5.19	6.56
	AA+	2.28	5.96	6.19
	AA-	2.84	7.69	10.53
	A	5.42	9.76	7.31
	BBB+	8.30	11.79	12.12
Chengtou Bonds	AAA	2.12	5.27	6.35
	AA+	2.61	6.00	6.04
	AA	2.96	6.67	6.16
	AA-	3.18	7.04	7.61
Company Bonds	AAA	2.45	5.89	8.33
	AA+	2.96	8.10	10.42
	AA	3.11	7.04	9.64
	AA-	4.14	7.15	9.30
avarage		3.50	7.31	7.95

Table 1: Fitting RMSE with the whole sample under different approaches, averaged over maturities. All numbers are of order 10^{-2}.

In-Sample Fitting - Rolling Windows

		MFMR	DNS	PCA
Enterprise Bonds	AAA	1.53	4.92	4.95
	AA+	1.91	5.91	5.05
	AA-	2.80	8.01	5.88
	A	5.08	9.71	8.05
	BBB+	7.51	11.77	12.62
Chengtou Bonds	AAA	1.66	5.10	5.10
	AA+	2.08	6.09	5.82
	AA	2.37	6.94	5.19
	AA-	2.68	7.33	5.99
Company Bonds	AAA	1.92	6.10	5.70
	AA+	2.38	8.90	6.34
	AA	2.49	7.58	5.96
	AA-	3.63	6.58	8.57
avarage		3.03	7.40	6.47

Table 2: In-Sample RMSE with rolling windows (fixed width = 300), averaged over maturities. All numbers are of order 10^{-2}.

Out-of-Sample Forecasting - Rolling Windows

		MFMR	DNS	PCA
Enterprise Bonds	AAA	3.26	5.83	8.07
	AA+	3.33	6.67	8.45
	AA-	3.50	9.15	9.98
	A	3.61	10.59	16.52
	BBB+	3.80	12.99	26.44
Chengtou Bonds	AAA	3.29	6.01	8.75
	AA+	3.42	6.79	10.88
	AA	3.43	7.68	9.42
	AA-	3.49	8.12	10.98
Company Bonds	AAA	3.82	7.09	9.04
	AA+	4.15	9.28	9.95
	AA	4.01	8.58	9.98
	AA-	4.12	8.29	17.85
avarage		3.59	8.30	11.94

Table 3: Out-of-Sample RMSE with rolling windows (fixed width = 300, one step ahead), averaged over maturities. All numbers are of order 10^{-2}.

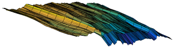
Concluding Remarks

- jointly modelling high-dimensional spread curves
- multivariate factorisable regression with high-dimensional functional data
- latent risky factors - systemic liquidity and dispersion measure
- robust forecasting outperforms DNS

Jointly Modelling and Robust Forecasting High-Dimensional Yield Curves

Wolfgang K. Härdle
Chen Huang
Linlin Niu

Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin
Wang Yanan Institute for Studies in Economics
Xiamen University
http://lvb.wiwi.hu-berlin.de
http://irtg1792.hu-berlin.de
http://wise.xmu.edu.cn/english



References

Chao, S.-K., Härdle, W.K. and Yuan, M. (2015)
Factorisable Sparse Tail Event Curves
SFB 649 Discussion Paper 2015-034, 2015

Chao, S.-K., Härdle, W.K. and Huang, C. (2016)
Multivariate Factorisable Sparse Asymmetric Least Squares Regression
SFB 649 Discussion Paper 2016-058, 2016

References

Diebold, F.X. and Li, C. (2006)
Forecasting the Term Structure of Government Bond Yields
Journal of Econometrics, 130, 337-364

Diebold, F.X. and Rudebusch, G.D. (2013)
Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach
Princeton University Press