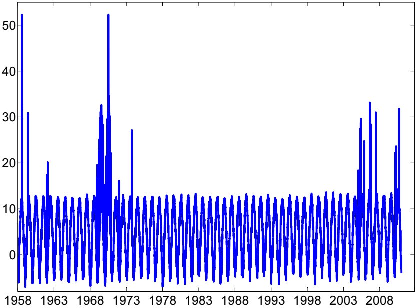
High Dimensional Nonstationary Time Series Modeling

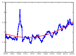
Song Song
Wolfgang Karl Härdle
Ya'acov Ritov

Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
University of California, Berkeley
The Hebrew University of Jerusalem

Temperatures and Climate Change

- China Meteorological Administration, $J = 159$ weather stations in China from 19570101 - 20091231
- Daily observations (averaged over stations), $T = 19358$





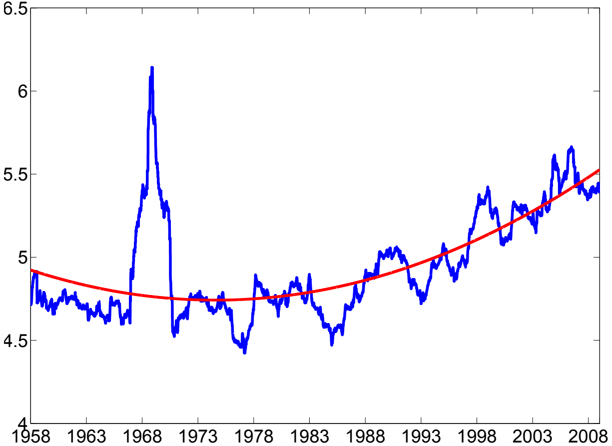


Figure 1: The moving average (of 730 nearby days) temperatures of China of 19570101 - 20091231

Weather Derivatives

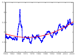
Econometrics

$$T_t$$
$$\downarrow$$
$$X_t = T_t - \Lambda_t$$
$$\downarrow$$
$$X_{t+3} = a^\top X_t + \sigma_t \varepsilon_t$$
$$\downarrow$$
$$\hat{\varepsilon}_t = \frac{\hat{X}_t}{\hat{\sigma}_t} \sim N(0,1)$$

Fin. Mathematics

$$CAR(3)$$
$$\downarrow$$
$$F_{CAT(t,\tau_1,\tau_2)}$$
$$\downarrow$$
$$MPR$$

- Detect complex trends, evaluate "non priced" places



Risk Perception

functional Magnetic Resonance Imaging

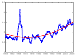
measures the oxygen level (BOLD) in the blood every 2-3 sec

Data Set

Series of 3-dim images

- each scan transformed on the resolution $2 \times 2 \times 2mm^3$
- 91 slices
- observed every 2.5 seconds
- data set: series of $T = 1360$ images with $91 \times 109 \times 91$ voxels

High-dimensional, high frequency & large data set.



Functional Magnetic Resonance Imaging

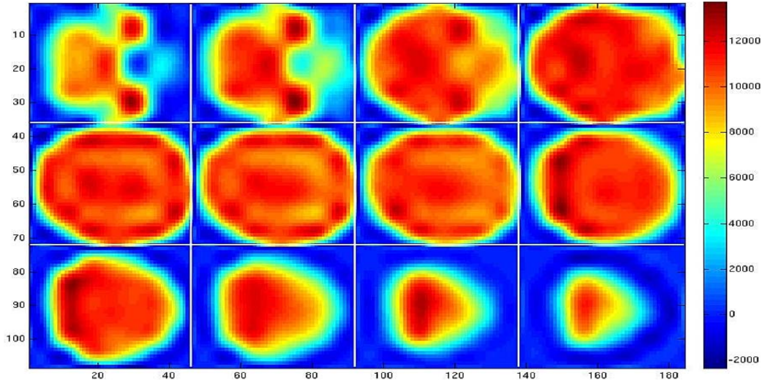
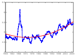


Figure 2: Typical example of fMRI image in a particular time point, 12 different horizontal slices of the brain's scan. fMRI

Risk Patterns and Brain Activities

- Which part is activated during *risk related decisions* ?
- Can statistical analysis help to detect this area?
- Can we provide an *integrated* dynamic analysis?
- Response curve (to stimuli)? classify "risky people"?



Implied Volatility Surface

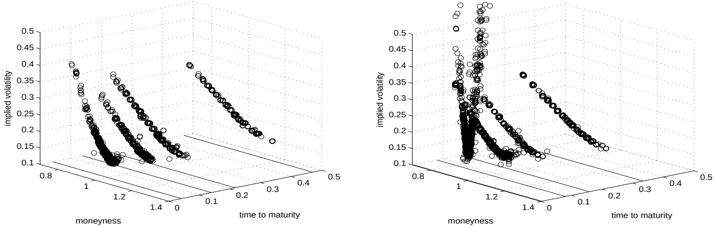
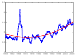


Figure 3: *Typical IV data design on two different days. Bottom solid lines indicate the observed maturities, which move towards the expiry. Left panel: observations on 20040701. Right panel: observations on 20040819.*

IVS

IVS

- "Strings", "Smile", "Skewness", Dimensionality not fixed
- Trading, hedging and risk management of option portfolios
- IVS reflects perception of market risk, Bakshi et al. (2000)...



Order Book

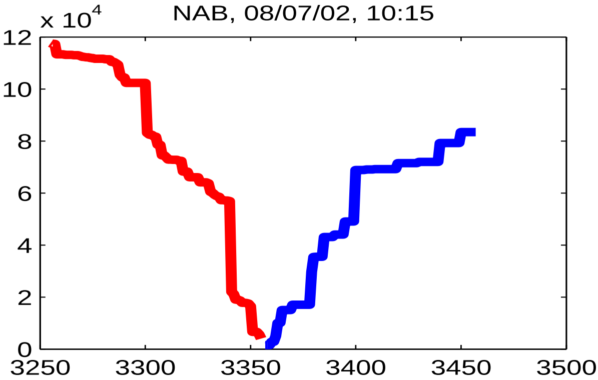


Figure 4: Bid and ask curves constructed from the order book of National Australian Bank stock prices on 20020801. LOB

Collateralized Debt Obligation

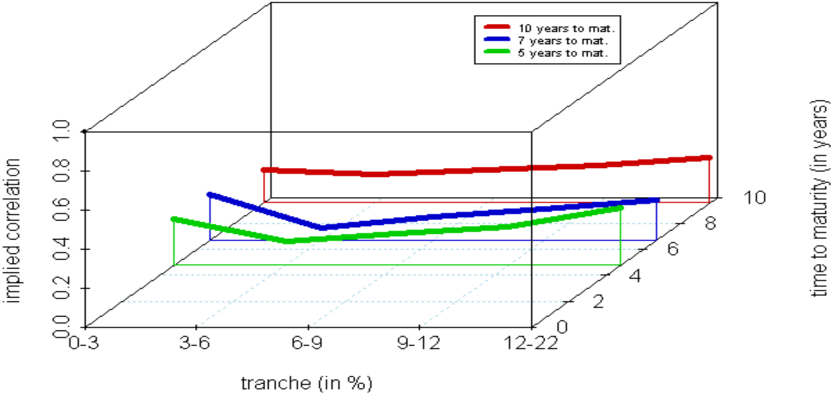
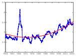


Figure 5: Compound correlations on 070321 w.r.t. time to maturity (in years), implied correlation and tranche (in %).  CDO



CO_2 Emission Allowance

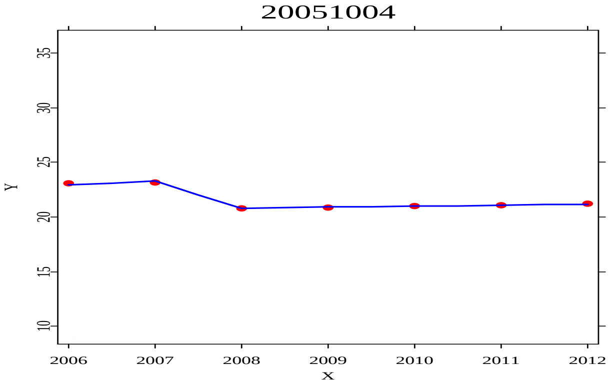
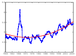


Figure 6: Term structure for CO_2 emission allowance's spot and futures prices, trading on 20051004 in the EEX market. CO2



Empirical Pricing Kernel

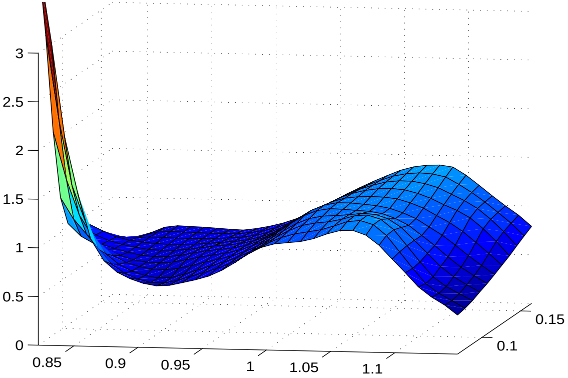


Figure 7: Estimated PK across moneyness κ and maturity τ at t = 20010710. EPK

Electricity Forward Prices

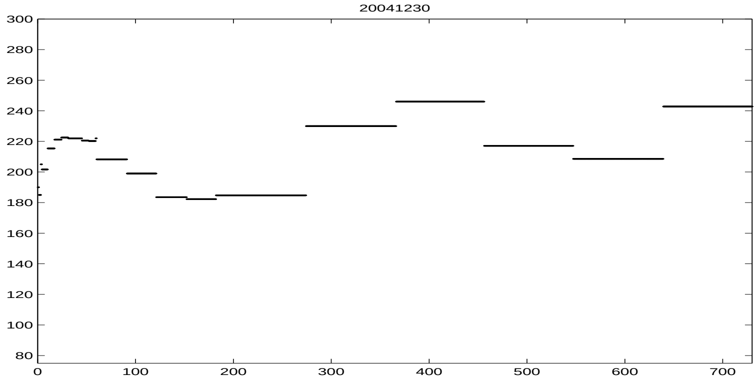


Figure 8: Term structure of the electricity prices (NOK/MWh) from the Nord Pool on 20041230. EFP

High Dimensional Nonstationary Time Series

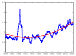
- Medicine
  - fMRI data
- Meteorology
  - Temperatures, rainfall etc. from many stations
- Finance
  - Implied Volatility Surface
  - Order Book
  - Collateralized Debt Obligation
  - CO_2 Emission Allowance
  - Empirical Pricing Kernel
  - Electricity Forward Price
- “eBird”, tree rings, . . .

Formal Setting

$$\underbrace{(X_{1,1}, Y_{1,1}), \ldots, (X_{J_1,1}, Y_{J_1,1})}_{t=1} \underbrace{(X_{1,2}, Y_{1,2}), \ldots}_{t=2} \ldots \underbrace{\ldots (X_{J_T,T}, Y_{1,T})}_{t=T}$$

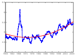
where:
$X_{j,t} \in \mathbb{R}^d, \; Y_{j,t} \in \mathbb{R}$
T - the number of observed time periods (days)
J_t - the number of the observations in (day) t
$\mathsf{E}(Y_t | X_t) = F_t(X_t)$.

What is $F_t(X_t)$? How it moves?



Basic Idea

- Use a "time & space" dynamic approach
- Separate time dynamics from space functions
- Low dim time series dynamics
- High dim (time invariant) space functions
- $\#$ of factors $\nearrow J$, fitting $\nearrow$



Dynamic Semiparametric Factor Model

$$E(Y_t|X_t) = \sum_{l=1}^{L} Z_{0,t,l} m_l(X_t) = Z_{0,t}^\top m(X_t) = Z_{0,t}^\top A\psi(X)$$

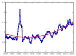
- $Z_t = (Z_{0,t,1}, \dots, Z_{0,t,L})^\top$ low dim (stationary) time series
- $m(\cdot)$ tuple of functions $(m_0, m_1, \dots, m_L)^\top$
- $\psi(x) = (\psi_1, \dots, \psi_K)^\top(x)$ vector of known basis functions
- A: $L \times K$ coefficient matrix.

Deterministic Model

$$Y_{t,j} = \sum_{l=1}^{L}\sum_{r=1}^{R} u_r(t)\gamma_{rl}\sum_{k=1}^{K} a_{lk}\psi_k(X_{t,j}) + \varepsilon_{tj} \quad (1)$$

$$Y_t^\top = \underbrace{U_t^\top \Gamma^*}_{Z_t^\top}\underbrace{A^*\Psi_t}_{m} + \varepsilon_t \stackrel{\text{def}}{=} U_t^\top \beta^{*\top}\Psi_t + \varepsilon_t. \quad (2)$$

- $U_t^\top = (u_1(t), \ldots, u_R(t))$, $u_r(t)$ time basis
- $\Psi_t = (\psi_1(X_t), \ldots, \psi_K(X_t))^\top$, $\psi_k(x)$ space basis
- $\beta^{*\top}$ $R \times K$ matrix consisting of β_{rk}s
  $\|\beta\|_{2,1} = \sum_{r=1}^{R}\sqrt{\sum_{k=1}^{K}\beta_{rk}^2}$ (group Lasso)

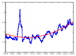


Generalized Dynamic Semiparametric Factor Model

$$
\begin{aligned}
Y_t^\top &= (Z_{0,t}^\top + U_t^\top \Gamma) A \Psi_t + \varepsilon_t' = U_t^\top \Gamma A \Psi_t + (Z_{0,t}^\top A \Psi_t + \varepsilon_t') \\
&\stackrel{\text{def}}{=} U_t^\top \Gamma A \Psi_t + \varepsilon_t, \quad \textit{with} \quad \mathsf{E}(Z_{0,t} | X_t) = 0.
\end{aligned}
$$

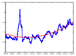
2 Step Estimation Procedure

- Find the trend based on $Y_t^\top = U_t^\top \Gamma A \Psi_t + \varepsilon_t$
- Based on $\widehat{\widehat{Y}}_t^\top \stackrel{\text{def}}{=} Y_t^\top - U_t^\top \widehat{\beta} \Psi_t$, $\widehat{A}$ and Ψ_t, obtain $\widehat{Z}_{0,t}$



Questions

- How to fit GDSFM?
- What risk is involved?
- How to select basis?

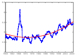


Overview

1. Data & Motivation✓
2. Estimation
3. Its Properties
4. Generalized Dynamic Semiparametric Factor Model
5. Weather, fMRI and IVS
6. How well are we doing?

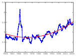
Time Basis

- Global trend: 1, t, $(3t^2 - 1)/2, \dots$ Legendre Polynomial
- Local variation: $\sin\{it/(p2\pi)\}$, $\cos\{it/(p2\pi)\}$, $i = 1, \dots$ Fourier Series
- Period $p = 11.8$ (fMRI), $365, 3650$ (weather)



Lasso & Group Lasso

- Lasso (least absolute shrinkage and selection operator), Tibshirani (1996)
- Group Lasso: Yuan and Lin (2006)
- Shrink some coefficients to 0
- Advantages over subset and ridge regressions



Space Basis

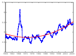
- FPCA (data dependent basis), motivated by Hall et. al (2006)
- ψ_k: eigenfunctions of the smoothed covariance operator

$$\widehat{\psi}(u,v) = \widehat{a}_0(u,v) - \widehat{a}(u)\widehat{a}(v)$$

$$\sum_{t=1}^{T}\sum_{j=1}^{J_t}\{Y_{tj} - a - \sum_{c=1}^{d} b^c(u^c - X_{tj}^c)\}^2 K\left(\frac{X_{tj}-u}{h_\mu}\right)$$

$$\sum_{t=1}^{T}\sum_{1\leqslant j\neq k\leqslant J_t}\{Y_{tj}Y_{tk} - a_0 - \sum_{c=1}^{d} b_1^c(u^c - X_{tj}^c) - \sum_{c=1}^{d} b_2^c(v^c - X_{tk}^c)\}^2$$

$$\times K\left(\frac{X_{tj}-u}{h_\phi}\right)K\left(\frac{X_{tj}-v}{h_\phi}\right).$$



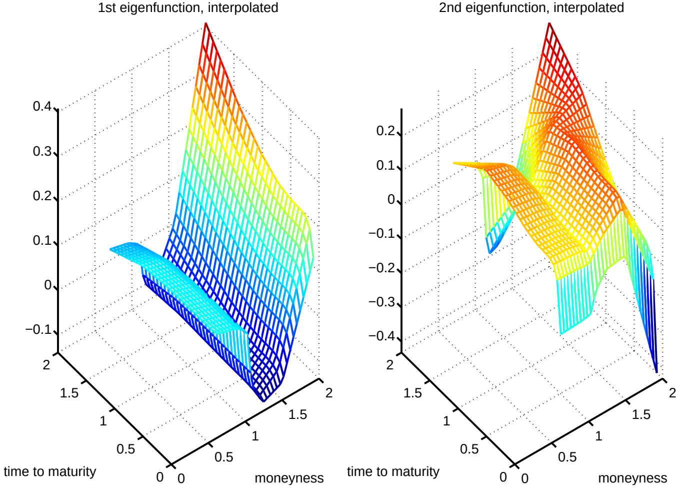
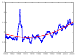


Figure 9: IVS Modeling using FPCA for space basis.



Estimation Procedure

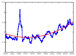
0 Space basis Ψ_t via FPCA

1 Time basis U_t selection via group Lasso; $\widehat{\beta}$ $(R \times K)$

$$\min_{\beta} (TJ)^{-1} \sum_{t=1}^{T} \left(Y_t^\top - U_t^\top \beta^\top \Psi_t \right) \left(Y_t^\top - U_t^\top \beta^\top \Psi_t \right)^\top + 2\lambda \|\beta\|_{2,1} \quad (3)$$

2 Split $\widehat{\beta}$ into $\widehat{\Gamma}$, $\widehat{A}$

$\widehat{\Gamma}$: first "L" eigenvectors of $\widehat{\beta}\widehat{\beta}^\top$; $\quad \widehat{A} = \widehat{\Gamma}^\top \widehat{\beta}$

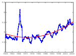


Tuning parameter λ

- Take 100 equally spaced $\lambda \in [0, \max_r \| \sum_t \Psi_t Y_t U_{tr} \| / \sqrt{K}]$
- Evaluate $C_p(\lambda)$ where

$$
\begin{aligned}
C_p(\lambda) &= \frac{\sum_t \| Y_t^\top - U_t^\top \widehat{\beta}^\top \Psi_t \|^2}{\tilde{\sigma}^2} - JT + 2df \\
\tilde{\sigma}^2 &= \frac{\sum_t \| Y_t^\top - U_t^\top \widehat{\beta}_{OLS}^\top \Psi_t \|^2}{JT - df} \\
df &= \sum_r \mathbf{1}\{\| \widehat{\beta}_r \| > 0\} + \sum_r \frac{\| \widehat{\beta}_r \|}{\| \widehat{\beta}^{OLS} \|}(K-1)
\end{aligned}
$$

- Choose the minimal $C_p(\lambda)$



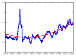
Theorem 1: Risk Bound (Gaussian)

Under Assumptions (A1, A2, A3, A4, 1), let $\lambda = 2\sigma/\sqrt{JT}\sqrt{1 + A\log R/\sqrt{T}}$ with $A > 8$, and then with probability at least $1 - R^{1-q}$ with $q = \min(A\log R, \sqrt{T})$, for all $\widehat{\beta}$:

$$(JT)^{-1}\sum_{t=1}^{T} \| \Psi_t^\top(\widehat{\beta} - \beta^*)U_t \|^2 \leqslant 64\sigma^2 s(1 + A\log R/\sqrt{T})/(\kappa^2 J),$$

$$T^{-1/2}\| \widehat{\beta} - \beta^* \|_{2,1} \leqslant 32\sigma s\sqrt{1 + A\log R/\sqrt{T}}/(\kappa^2\sqrt{J}),$$

$$M(\widehat{\beta}) \leqslant 64\phi^2_{max} s/\kappa^2$$



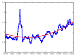
Theorem 2: Risk Bound (Non-Gaussian)

Under Assumptions (A1, A2, A3, A5, 1, 2), let $\lambda = \sigma\sqrt{(\log R)^{1+\delta}/JT}$, $\delta > 0$, and then with probability at least $1 - (2e\log R - e)C/(\log R)^{1+\delta}$, for all $\widehat{\beta}$:

$$(JT)^{-1}\sum_{t=1}^{T} \| \Psi_t^\top(\widehat{\beta} - \beta^*)U_t \|^2 \leqslant 16\sigma^2 s(\log R)^{1+\delta}/(\kappa^2 J)$$

$$T^{-1/2}\| \widehat{\beta} - \beta^* \|_{2,1} \leqslant 16\sigma s\sqrt{(\log R)^{1+\delta}}/(\kappa^2\sqrt{J})$$

$$M(\widehat{\beta}) \leqslant 64\phi_{max}^2 s/\kappa^2$$



Theorem 3: Risk Bound (Dependent)

Under Assumptions (A1, A2, A3, 1, 3), let

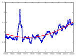
$$\lambda = \frac{C'}{\sqrt{T}} + \sqrt{\frac{\mathcal{X}^*(\mathcal{T}) \sum_t b_t^2}{(\log R)^{1-\delta'} T^2}}, \quad \delta' > 0,$$

and then with probability at least $p(1 - R^{-\delta'})$, for $\forall\ \widehat{\beta}$:

$$(JT)^{-1} \sum_{t=1}^{T} \| \Psi_t^\top (\widehat{\beta} - \beta^*) U_t \|^2 \leqslant 16 \left(C' + \sqrt{\frac{\mathcal{X}^*(\mathcal{T}) \sum_t b_t^2}{(\log R)^{1-\delta'} T}} \right)^2 s/\kappa^2$$

$$T^{-1/2} \| \widehat{\beta} - \beta^* \|_{2,1} \leqslant 16 \left(C' + \sqrt{\frac{\mathcal{X}^*(\mathcal{T}) \sum_t b_t^2}{(\log R)^{1-\delta'} T}} \right) s/\kappa^2$$

$$M(\widehat{\beta}) \leqslant 64 \phi_{max}^2 s/\kappa^2$$



Comparison

i.i.d. Gaussian

- Dependence on R is negligible for large T
- Low sparsity, s/κ^2 large, bounds large, $M(\widehat{\beta})$ large

Independent, bounded 2nd moment

- Dependence on R not made negligible for large T

Dependent

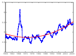
- Dependence level $\nearrow$, bound $\nearrow$

Theorem 4: $\widehat{Z}_{0,t}$ not get affected

Suppose that all assumptions in Theorem 3 and Assumptions (B1 - B7) hold. Then we have

$$\frac{1}{T}\sum_{1\leq t\leq T}\left\|\widehat{Z}_{0,t}^{\top}\widehat{A}-Z_{0,t}^{\top}A^{*}\right\|^{2}=\mathcal{O}_{P}(\rho^{2}+\delta_{K}^{2}). \quad (4)$$

since $\widehat{\beta}$ is close enough to β



Definitions

$$B \stackrel{\text{def}}{=} (\sum_{t=1}^{T} Z_{0,t}\widehat{Z}_{0,t})^{-1} \sum_{t=1}^{T} Z_{0,t} Z_{0,t}^{\top}$$

$$\widetilde{Z}_{0,t} \stackrel{\text{def}}{=} B^{\top}\widehat{Z}_{0,t}$$

$$\widetilde{Z}_{n,t} \stackrel{\text{def}}{=} (T^{-1}\sum_{s=1}^{T} \widetilde{Z}_{0,s}\widetilde{Z}_{0,s}^{\top})^{-1/2}\widetilde{Z}_{0,t}$$

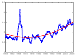
$$Z_{n,t} \stackrel{\text{def}}{=} (T^{-1}\sum_{s=1}^{T} Z_{0,s}Z_{0,s}^{\top})^{-1/2} Z_{0,t}$$

Theorem 5: Covariance Equivalence

Suppose that all assumptions in Theorem 3 and Assumptions (B1 - B7, C1 - C2) hold. Then we have for $h \geq 0$

$$T^{-1} \sum_{t=\max[1,-h+1]}^{\min[T,T-h]} \tilde{Z}_{0,t} \left(\tilde{Z}_{0,t+h} - \tilde{Z}_{0,t} \right)^\top - Z_{0,t} \left(Z_{0,t+h} - Z_{0,t} \right)^\top = \mathcal{O}_P(T^{-1/2})$$

$$T^{-1} \sum_{t=\max[1,-h+1]}^{\min[T,T-h]} \tilde{Z}_{n,t} \tilde{Z}_{n,t+h}^\top - Z_{n,t} Z_{n,t+h}^\top = \mathcal{O}_P(T^{-1/2})$$



Weather: space functions

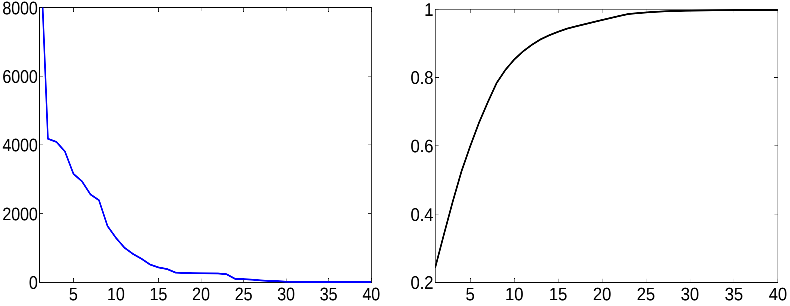
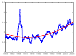
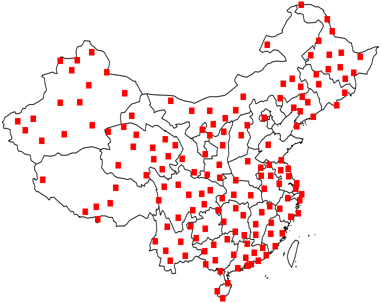


Figure 10: Distribution of the eigenvalues and the relative proportion of variance explained by the first K basis.



China Climate Types



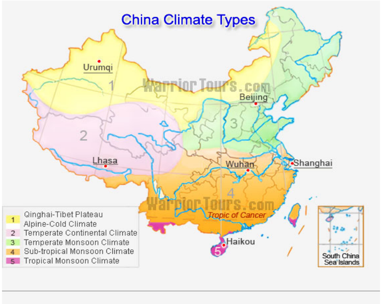
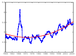


Figure 11: Weather stations & China Climate Types

Time Basis

	Factors		Factors
Trend	1	Large	$\sin\{2\pi t/(365 \cdot 10)\}$
(Year by Year)	t	Period	$\cos\{2\pi t/(365 \cdot 10)\}$
	$3t^2 - 1$		$\sin\{4\pi t/(365 \cdot 10)\}$
Seasonal	$\sin\{2\pi t/365\}$		$\cos\{4\pi t/(365 \cdot 10)\}$
Effect	$\cos\{2\pi t/365\}$		$\sin\{6\pi t/(365 \cdot 10)\}$
	$\ldots$		$\ldots$
	$\cos\{20\pi t/365\}$		$\cos\{20\pi t/(365 \cdot 10)\}$

Table 1: Initial choice of $53 \cdot 3 + 20 = 179$ time basis.



Time Basis Coefficients

- Long term: quadratic trend, warming effect
- Yearly variation ($p = 365$): earth rotation
- 10-year variation ($p = 3650$): solar activity

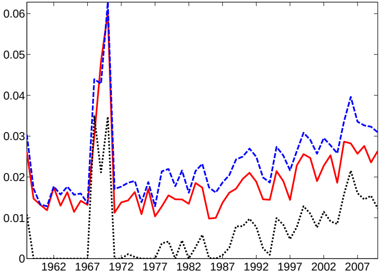
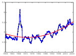


Figure 12: Estimated coefficients of the 1st factor $\widehat{\Gamma}_{r1}$ w.r.t. the yearly polynomial time basis (constant, linear, quadratic).



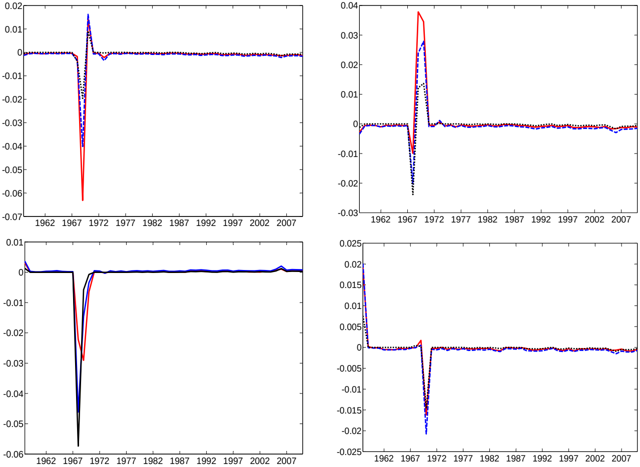
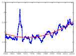
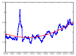


Figure 13: Estimated coefficients of the 2nd - 5th factors $\widehat{\Gamma}_{r2}, \ldots, \widehat{\Gamma}_{r5}$ w.r.t. the $54 \cdot 3$ yearly polynomial time basis.



Basis			Estimates		
$\sin 2\pi t/365$	-0.1777	0.0076	0.0177	-0.0136	0.0084
$\cos 2\pi t/365$	-0.6081	0.0126	0.0366	-0.0369	0.0114
$\sin 4\pi t/365$	0.0000	0.0000	0.0000	0.0000	0.0000
$\cos 4\pi t/365$	-0.0145	0.0028	0.0021	-0.0022	0.0029
...	0.0000	...			
$\cos 20\pi t/365$	0.0000	...			
$\sin 2\pi t/(365 \cdot 10)$	0.0025	-0.0006	0.0009	-0.0008	-0.0001
$\cos 2\pi t/(365 \cdot 10)$	0.0000	...			
...	0.0000	...			
$\cos 20\pi t/(365 \cdot 10)$	0.0000	...			

Table 2: Estimated coefficients of the 5 factors w.r.t. the 20 Fourier series time basis.



Extracted Trends

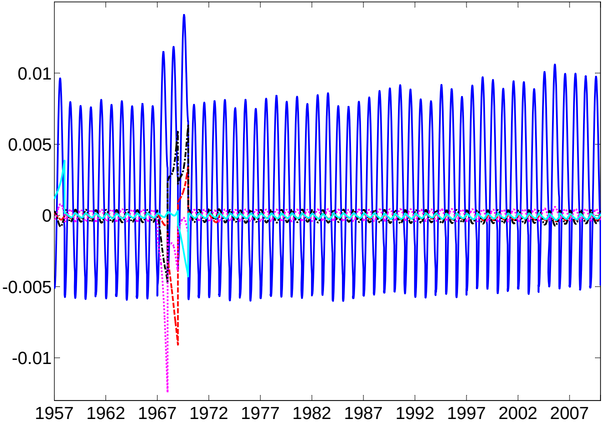
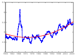


Figure 14: Extracted trends based on $U_t^\top \widehat{\beta}$.



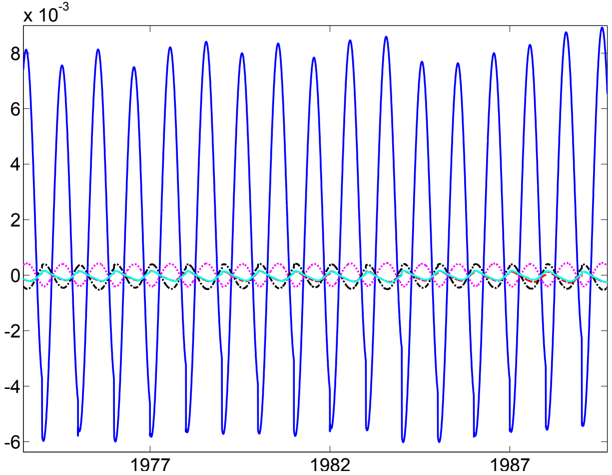
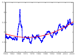


Figure 15: Extracted trends based on $U_t^\top \widehat{\beta}$.



Estimated Stochastic Process $\widehat{Z}_{0,t,1}$

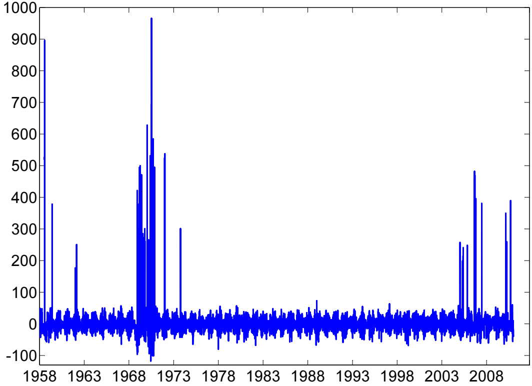
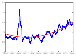


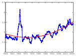
Figure 16: Estimated Stochastic Process $\widehat{Z}_{0,t,1}$



Econometric Modeling of $\widehat{Z}_{0,t}$

$\widehat{Z}_{0,t} = \mathcal{R}\widehat{Z}_{0,t-1} + \varepsilon_{0,t}$ with random vector $\varepsilon_{0,t}$ and estimated coefficient matrix:

$$\begin{pmatrix} 0.9732 & -0.0135 & -0.0002 & -0.0006 & -0.0002 \\ 0.0127 & 0.1766 & -0.1824 & -0.0682 & -0.0009 \\ 0.0358 & -0.2867 & 0.4493 & -0.1138 & 0.0053 \\ -0.0001 & -0.1967 & -0.1962 & 0.8010 & -0.0052 \\ 0.0790 & 0.0492 & 0.0690 & -0.0225 & 0.8418 \end{pmatrix}.$$



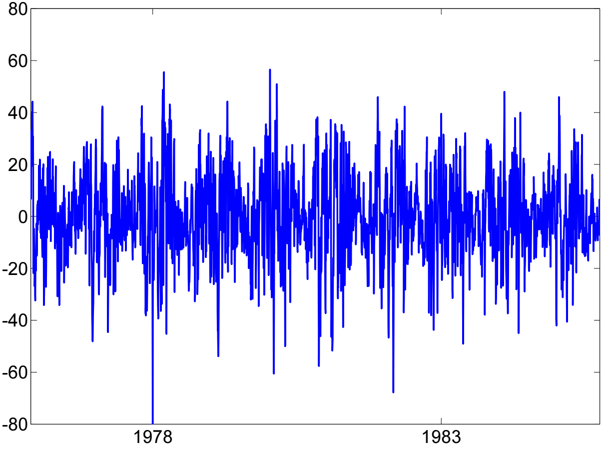


Figure 17: Estimated Stochastic Process $\widehat{Z}_{0,t,1}$

Improvements

- Periodic behavior in $\widehat{Z}_{0,t}$
- Step 1 (detrending, both weather + fMRI):
  Variance of the noise $\sim$ mean of the measurements!
- Group Lasso under heteroscedasticity with high J (Poisson - like model)

$$Y_t^\top = U_t^\top \Gamma A \Psi_t + \varepsilon_t, \quad \text{Cov}(\varepsilon_t) = diag(|U_t^\top \Gamma A \Psi_t|)$$

- Lasso under heteroscedasticity, Jia et. al. (2009)

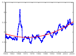
Current calibration

- 2 Steps: Fourier truncated series + GARCH(p,q) $\hat{\sigma}^2_{t,FTSG}$

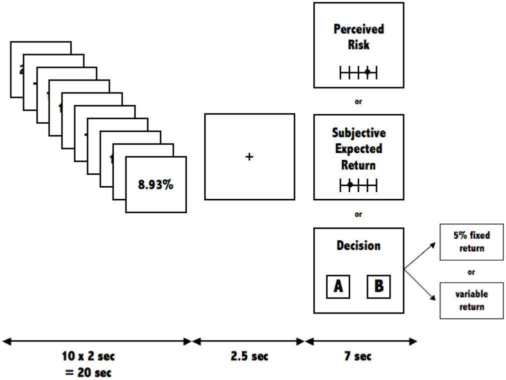
$$\begin{aligned}\sigma_t^2 &= c_1 + \sum_{i=1}^{16}\left\{c_{2i}\cos\left(\frac{2i\pi t}{365}\right) + c_{2i+1}\sin\left(\frac{2i\pi t}{365}\right)\right\}\\ &+ \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2\end{aligned}$$

- 1 Step: Local linear Regression (LLR) $\hat{\sigma}^2_{t,LLR}$, $Y_i = \hat{\varepsilon}^2_{t_i}$

$$\min_{a,b}\sum_{i=1}^{n}\left\{Y_i - a(t) - b(t)(t_i - t)\right\}^2 K\left(\frac{t_i - t}{h}\right)$$



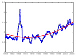
Risk Perception and Investment Decision



Returns Pause Decision

fMRI methods (Panel)

- existing methods to analyze these data: voxel-wise GLM
  - strong a priori hypothesis necessary
- new statistical method: DSFM
  - dimension reduction keeping the data structure
  - exploratory analysis



Panel Version Model with Multi Subjects $1 \leq i \leq I$

$$Y^i_{t,j} = \sum_{l=1}^{L} (\alpha^i_{t,l} + U_t^\top \Gamma^i_l) m_l(X_{t,j}) + \varepsilon_{t,j}, \;\; 1 \leq j \leq J_t, \; 1 \leq t \leq T,$$

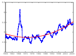
with fixed effect $\alpha^i_{t,l}$ and

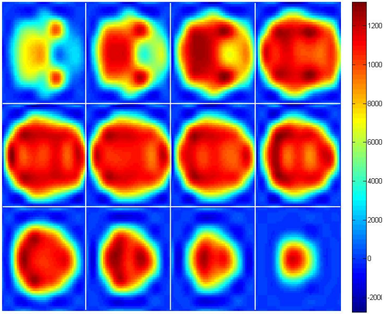
$$\sum_{i=1}^{I} \left(\sum_{l=1}^{L} \alpha^i_{t,l} m_l(X_{t,j}) | X_{t,j} \right) = 0.$$

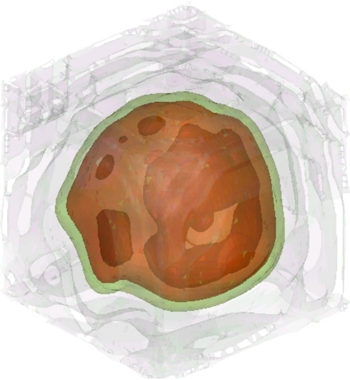
2-step estimation procedure

1. Average $Y^i_{t,j}$ over i, and estimate common m_l as in the original approach.
2. Given the common m_l, for i, estimate their specific factors in time $Z^i_{t,l}$.

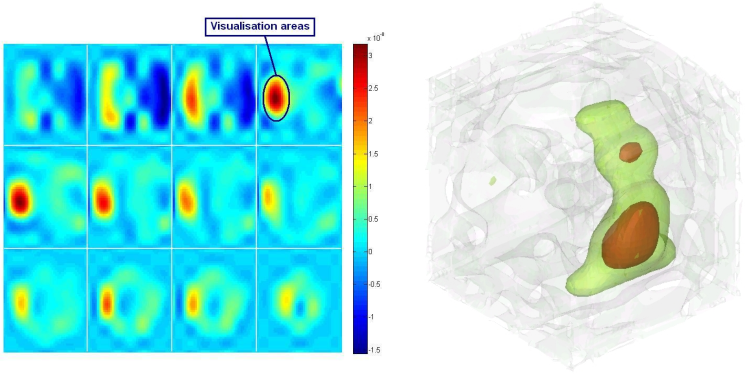
$$Y^i_{t,j} = \sum_{l=1}^{L} U_t^\top \Gamma^i_l \overline{m}_l(X_{t,j}) + \varepsilon^i_{t,j}$$



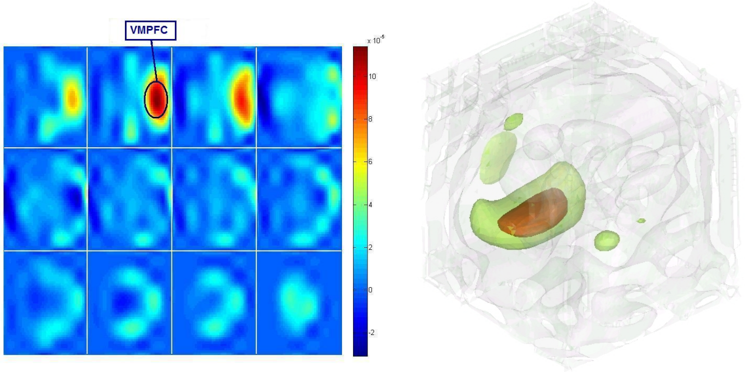




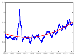
Estimated factor loading $\hat{m}_1$ with $L = 5$.

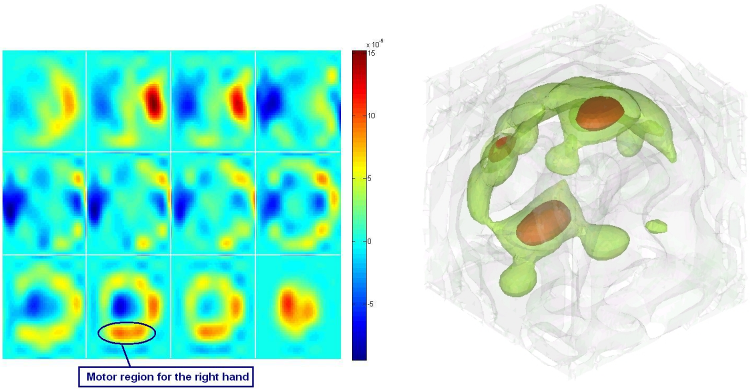


Estimated factor loading $\hat{m}_2$ with $L = 5$.

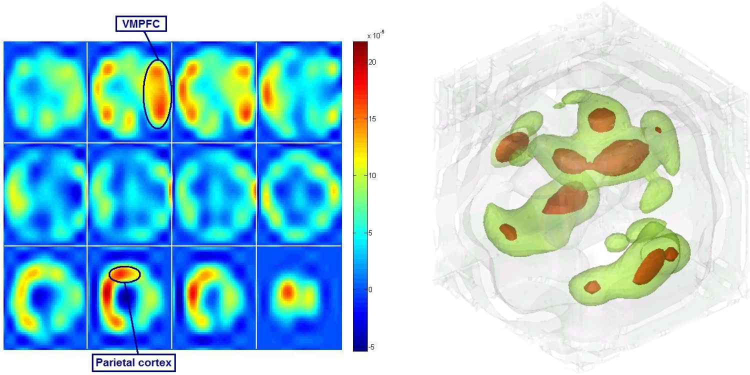


Estimated factor loading $\hat{m}_3$ with $L = 5$.
(VMPFC = Ventromedial prefrontal cortex)





Estimated factor loading $\hat{m}_4$ with $L = 5$.



Estimated factor loading $\hat{m}_5$ with $L = 5$.

Response to Stimuli

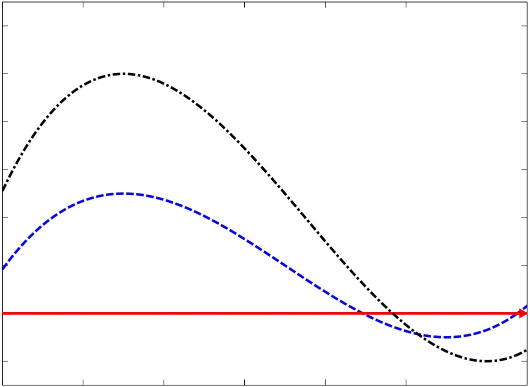
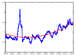


Figure 18: Response curves (to stimuli) $U_t^\top \widehat{\Gamma}_2^i$ for probands $i = 9$ & $i = 19$ with periodic cubic polynomial as time basis.



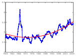
SVM Analysis (Risk)

- Different subjects' response curves have different shapes
- SVM based on the $\widehat{\beta}$

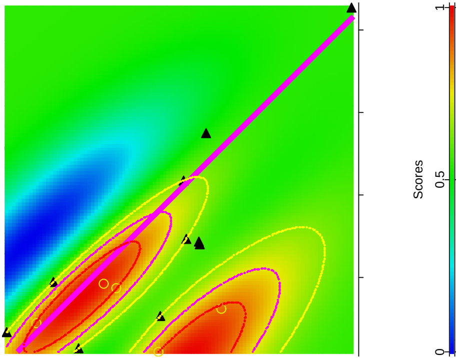
MEAN		Estimated	
Data	Strongly	0.85	0.14
	Weakly	0.59	0.40

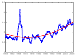
Table 3: Classification rates of the SVM method.

The rates hold over a wide range of parameters



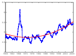
SVM Classification





Related & Future Research

- Risk Patterns and Correlated Brain Activities (with A. Myšičková)
- Tell what the subject see
- Emotions recognition from face (FOX: “Don’t lie to me”) + fMRI analysis



Implied Volatility Surface Modeling

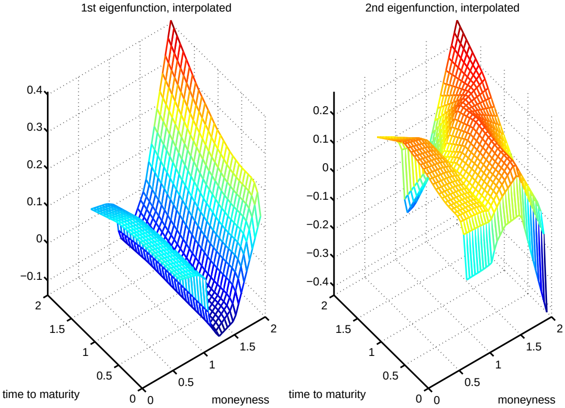
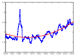


Figure 19: IVS Modeling using FPCA for space basis.



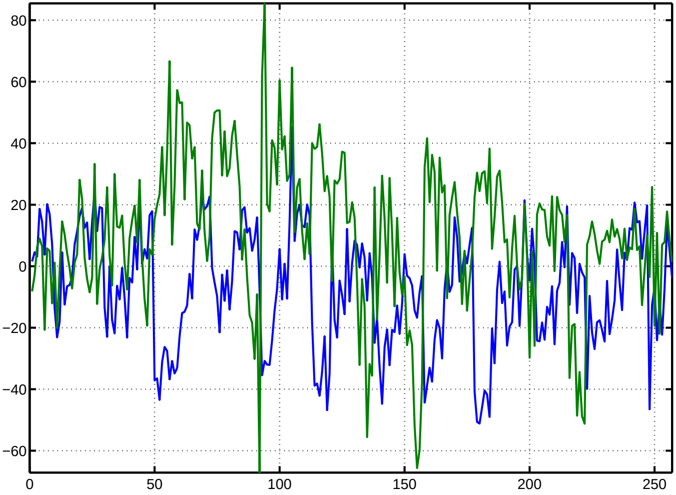


Figure 20: Estimated time series of factors $\widehat{Z}_{t,1}$, $\widehat{Z}_{t,2}$ (20070102 - 20081230)

High Dimensional Nonstationary Time Series Modeling

Song Song
Wolfgang Karl Härdle
Ya'acov Ritov

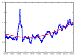
Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
University of California, Berkeley
The Hebrew University of Jerusalem

References

Bakshi, G., Cao, C. and Chen, Z.
"Do call and underlying prices always move in the same direction?"
Review of Financial Studies, 13(3): 549–584, 2000.

Fengler, M., Härdle, W. and Mammen, E.
A Dynamic Semiparametric Factor Model for Implied Volatility String Dynamics
Journal of Financial Econometrics, 5(2): 189–218, 2007.

Hall, P., Müller, H. G. and Wang J.
Properties of principal component methods for functional and longitudinal data analysis
Ann. Statist., 34(3): 1493–1517, 2006.

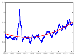


References

Janson, S.
Large deviations for sums of partly dependent random variables
Random Structures Algorithms, 24(3): 234–248, 2004.

Jia, J., Rohe, K. and Yu, B.
The Lasso under Heteroscadecity
Technical Report 783, Statistics Department, UC Berkeley., 2009.

Park, B., Mammen, E., Härdle, W. and Borak, S.
Time Series Modelling with Sempiparametric Factor Dynamics
J. Amer. Stat. Assoc., 104(485): 284-298, 2009.



References

Tibshirani, R.
Regression Shrinkage and Selection via the Lasso
J. R. Statist. Soc. B, 58: 267–288, 1996.

Assumptions

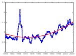
- A1 Normalization of U_t & Ψ_t: $\Psi_t \Psi_t^\top / J = I_K$, $\sum_{t=1}^{T} U_t^\top U_t / R = 1$
- A2 The number of nonzero β_r^*s: $M(\beta^*) \leq s$
- A3 ϕ_{max} is the maximum eigenvalue of $\sum_{t=1}^{T} U_t U_t^\top$
- A4 The error terms $\varepsilon_1, \ldots, \varepsilon_T$ are i.i.d. **Gaussian** with mean 0 and variance $\sigma^2 I_{J \times J}$
- A5 The error terms $\varepsilon_1, \ldots, \varepsilon_T$ are independent with mean 0 and **finite** variance $\mathrm{E}(\varepsilon_{tj}^2) \leqslant \sigma^2$

Assumption 1

There exists a positive number $\kappa = \kappa(s)$ *such that*

$$\min \left\{ \frac{\sum_t \|\Psi_t^\top \Delta U_t\|}{\sqrt{J} \parallel \Delta_{\mathcal{R}} \parallel} : |\mathcal{R}| \leqslant s, \Delta \in \mathbb{R}^{K \times R} \backslash \{0\}, \right.$$
$$\left. \parallel \Delta_{\mathcal{R}^c} \parallel_{2,1} \leqslant 3 \parallel \Delta_{\mathcal{R}} \parallel_{2,1} \right\} \geqslant \kappa.$$

- Restriction on the eigenvalues of U_t as a function of sparsity s
- Low sparsity, s big, κ small

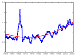


Assumption 2

The matrices Ψ_t and U_t are such that

$$(JT)^{-1} \sum_{t=1}^{T} \sum_{j=1}^{J} \left(\max_r \left| \sum_{k=1}^{K} \Psi_{tkj} U_{tr} \right| \right)^2 \leqslant C,$$

for a constant $C > 0$.

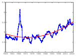


Measure of Dependence, Jason (2004)

Given a set $\mathcal{T}$ and random variables V_t, $t \in \mathcal{T}$, we say:

- A subset $\mathcal{T}'$ of $\mathcal{T}$ is *independent* if the corresponding random variables $\{V_t\}_{t\in\mathcal{T}'}$ are independent.
- A family $\{\mathcal{T}_j\}_j$ of subsets of $\mathcal{T}$ is a *cover* of $\mathcal{T}$ if $\bigcup_j \mathcal{T}_j = \mathcal{T}$.
- A family $\{(\mathcal{T}_j, w_j)\}_j$ of pairs $(\mathcal{T}_j, w_j)$, where $\mathcal{T}_j \subseteq \mathcal{T}$ and $w_j \in [0,1]$ is a *fractional cover* of $\mathcal{T}$ if $\sum_j w_j \mathbf{1}_{\mathcal{T}_j} \geqslant \mathbf{1}_{\mathcal{T}}$, i.e. $\sum_{j:t\in\mathcal{T}_j} w_j \geqslant 1$ for each $t \in \mathcal{T}$.
- A (fractional) cover is *proper* if each set $\mathcal{T}_j$ in it is independent.
- $\mathcal{X}(\mathcal{T})$ is the size of the smallest proper cover of $\mathcal{T}$, i.e. the smallest m such that $\mathcal{T}$ is the union of m independent subsets.
- $\mathcal{X}^*(\mathcal{T})$ is the minimum of $\sum_j w_j$ over all proper fractional covers $\{(\mathcal{T}_j, w_j)\}_j$.

$\mathcal{X}^*(\mathcal{T})$: measure of dependence; $\mathcal{X}^*(\mathcal{T}) = 1$ (independent).



Assumption 3

With a high probability p, Ψ_t, U_t *and* ε_t *are such that*

$$
\begin{aligned}
(J^{-1}\sum_{k=1}^{K}\sum_{j=1}^{J}\Psi_{tkj}\varepsilon_{tj}U_{tr})^2 &\leqslant b_t^2 \\
\mathrm{E}(JT)^{-1}\Big\{\sum_{t=1}^{T}(\sum_{k=1}^{K}\sum_{j=1}^{J}\Psi_{tkj}\varepsilon_{tj}U_{tr})^2\Big\}^{1/2} &\leqslant \frac{C'}{\sqrt{T}}.
\end{aligned}
$$

for $\forall\ r$ *and some constants* $b_t, C' > 0, t = 1, \ldots, T$.

Assumptions

B1 $X_{1,1}, \ldots, X_{T,J}$, $\varepsilon'_{1,1}, \ldots, \varepsilon'_{T,J}$, and $Z_{0,1}, ..., Z_{0,T}$ are independent.

B2 $X_{t,1}, \ldots, X_{t,J}$ are identically distributed, support $[0,1]^d$ and a density f_t that is bounded from below and above on $[0,1]^d$, uniformly over $t = 1, \ldots, T$.

B3 We assume that $\mathrm{E}\,\varepsilon'_{t,j} = 0$ for $1 \leq t \leq T, 1 \leq j \leq J$, and for $c > 0$ small enough $\sup_{1 \leq t \leq T, 1 \leq j \leq J} \mathrm{E} \exp\{c(\varepsilon'_{t,j})^2\} < \infty$.

B4 The vector of functions $m = (m_1, \ldots, m_L)^\top$ can be approximated by Ψ_k, i.e.

$$\delta_K \stackrel{\text{def}}{=} \sup_{x \in [0,1]^d} \inf_{A \in \mathbb{R}^{L \times K}} \|m(x) - A\Psi(x)\| \to 0$$

as $K \to \infty$. We denote A that fulfills $\sup_{x \in [0,1]^d} \|m(x) - A\Psi(x)\| \leq 2\delta_K$ by A^*.

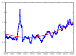
B5 There exist constants $0 < C_L < C_U < \infty$ such that all eigenvalues of the matrix $T^{-1}\sum_{t=1}^{T} Z_{0t}Z_{0,t}^{\top}$ lie in the interval $[C_L, C_U]$ with probability tending to one.

B6 The minimization (3) runs over all values β with

$$\sup_{x\in[0,1]^d} \max_{1\leqslant t\leqslant T} \|Z_{0,t}^{\top}A\Psi(x)\| \leqslant M_T,$$

where the constant M_T fulfils $\max_{1\leqslant t\leqslant T}\|Z_{0,t}\| \leqslant M_T/C_m$ (with probability tending to one) for a constant C_m such that $\sup_{x\in[0,1]^d}\|m(x)\| < C_m$.

B7 It holds that $\rho^2 = (K+T)M_T^2\log(JTM_T)/(JT) \to 0$. The dimension L is fixed.



Assumptions

C1 $Z_{0,t}$ is a strictly stationary sequence with $\mathrm{E}(Z_{0,t}) = 0$, $\mathrm{E}(\|Z_{0,t}\|^{\gamma}) < \infty$ for some $\gamma > 2$. It is strongly mixing with $\sum_{i=1}^{\infty} \alpha(i)^{(\gamma-2)/\gamma} < \infty$. The matrix $\mathrm{E}\, Z_{0,t} Z_{0,t}^{\top}$ has full rank. The process $Z_{0,t}$ is independent of $X_{11}, \ldots, X_{TJ}, \varepsilon'_{11}, \ldots, \varepsilon'_{TJ}$.

C2 It holds that

$$[\log(KT)^2\{(KM_T/J)^{1/2} + T^{1/2} M_T^4 J^{-2} + K^{3/2} J^{-1} + K^{4/3} J^{-2/3} T^{-1/6}\} + 1] T^{1/2}(\rho^2 + \delta_K^2) = \mathcal{O}(\rho^2 + \delta_K^2)$$