SIEVE ESTIMATION OF TIME-VARYING FACTOR LOADINGS

CHEUNG Ying Lun

December 28, 2017

Goethe University Frankfurt

- Standard factor model

$$X_{it} = \boldsymbol{\lambda}_i' \boldsymbol{F}_t + e_{it}$$

- X_{it}: observed variables, $i = 1, \ldots, N$, $t = 1, \ldots, T$
- $\boldsymbol{F}_t$: common factors, K-dimensional
- $\boldsymbol{\lambda}_i$: factor loadings

Introduction

- *Time-varying* factor model

$$X_{it} = \boldsymbol{\lambda}_{it}' \boldsymbol{F}_t + e_{it}$$

- X_{it}: observed variables, $i = 1, \ldots, N$, $t = 1, \ldots, T$
- $\boldsymbol{F}_t$: common factors, K-dimensional
- $\boldsymbol{\lambda}_{it}$: factor loadings

Introduction

- Standard factor model

$$X_{it} = \boldsymbol{\lambda}_i' \boldsymbol{F}_t + e_{it}$$

- *Time-varying* factor model

$$X_{it} = \boldsymbol{\lambda}_{it}' \boldsymbol{F}_t + e_{it}$$

Introduction

- Standard factor model

$$X_{it} = \boldsymbol{\lambda}_i' \mathbf{F}_t + e_{it}$$

  - NT data points
  - $NK + TK$ parameters

- *Time-varying* factor model

$$X_{it} = \boldsymbol{\lambda}_{it}' \mathbf{F}_t + e_{it}$$

Introduction

- Standard factor model

$$X_{it} = \boldsymbol{\lambda}_i' \mathbf{F}_t + e_{it}$$

  - NT data points
  - $NK + TK$ parameters

- *Time-varying* factor model

$$X_{it} = \boldsymbol{\lambda}_{it}' \mathbf{F}_t + e_{it}$$

  - NT data points
  - $NTK + TK$ parameters

Introduction

- Bates, Plagborg-Møller, Stock and Watson (2013):
  - Factor space can be consistently estimated by PCA if $\boldsymbol{\lambda}_{it}$ are
    - white noise sequences, or
    - random walks, scaled by a factor of $h = o(T^{-1/2})$
  - Estimation of $\boldsymbol{\lambda}_{it}$ not considered.

Introduction

- Bates, Plagborg-Møller, Stock and Watson (2013):
  - Factor space can be consistently estimated by PCA if $\boldsymbol{\lambda}_{it}$ are
    - white noise sequences, or
    - random walks, scaled by a factor of $h = o(T^{-1/2})$
  - Estimation of $\boldsymbol{\lambda}_{it}$ not considered.
- Su and Wang (2017):
  - $\boldsymbol{\lambda}_{it} = \boldsymbol{\lambda}(t/T)$ are some smooth functions
  - Estimate $\boldsymbol{F}_t$ and $\boldsymbol{\lambda}_{it}$ by a kernel-type local PCA
  - Estimators consistent up to a rotation matrix $\boldsymbol{H}^{(t)}$ for each t, i.e.,

$$\widehat{\boldsymbol{F}}_t \xrightarrow{p} \boldsymbol{H}^{(t)\prime}\boldsymbol{F}_t, \qquad \widehat{\boldsymbol{\lambda}}_{it} \xrightarrow{p} \boldsymbol{H}^{(t)-1}\boldsymbol{\lambda}_{it}.$$

Introduction

Questions to be asked (answered) in this paper:

1. Can we allow non-trivial variations in loadings and still estimate the factors *without rotation*?

Introduction

Questions to be asked (answered) in this paper:

1. Can we allow non-trivial variations in loadings and still estimate the factors *without rotation*?
2. Can we estimate the sample paths of all loadings?

Introduction

Questions to be asked (answered) in this paper:

1. Can we allow non-trivial variations in loadings and still estimate the factors *without rotation*?
2. Can we estimate the sample paths of all loadings?
3. Can we estimate them more efficiently if some of the loadings are constant?

Introduction

Questions to be asked (answered) in this paper:

1. Can we allow non-trivial variations in loadings and still estimate the factors *without rotation*?
2. Can we estimate the sample paths of all loadings?
3. Can we estimate them more efficiently if some of the loadings are constant?
4. Can we identify them if some of the loadings are constant?

Outlines

MODEL AND ASSUMPTIONS

Time-Varying Factor Model

$$X_{it} = \boldsymbol{\lambda}_{it}' \boldsymbol{F}_t + e_{it}, \qquad i = 1, \dots, N; t = 1, \dots, T.$$

- $\{\boldsymbol{F}_t\}$ is weakly stationary
- $\{e_{it}\}$ has limited time and cross-sectional dependence

Time-Varying Factor Model

$$X_{it} = \boldsymbol{\lambda}_{it}' \boldsymbol{F}_t + e_{it}, \qquad i = 1, \ldots, N; t = 1, \ldots, T.$$

- $\{\boldsymbol{F}_t\}$ is weakly stationary
- $\{e_{it}\}$ has limited time and cross-sectional dependence
- $\boldsymbol{\lambda}_{it} := \boldsymbol{\lambda}_i(t/T)$ is α-Hölder continuous
  - $\boldsymbol{\lambda}_i(\cdot)$ can be a smooth, differentiable function
  - $\boldsymbol{\lambda}_i(\cdot)$ can be a fractional Brownian motion

ESTIMATION PROCEDURE

Estimating Factors

- Consider for some t close to s,

$$\begin{aligned} X_{it} &= \boldsymbol{\lambda}_{it}' F_t + e_{it} \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{is})' F_t \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + w_{ist} \end{aligned}$$

Estimating Factors

- Consider for some t close to s,

$$\begin{aligned} X_{it} &= \boldsymbol{\lambda}_{it}' F_t + e_{it} \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{is})' F_t \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + w_{ist} \end{aligned}$$

- $w_{ist} \xrightarrow{p} 0$ as $(t - s)/T \to 0$

Estimating Factors

- Consider for some t close to s,

$$\begin{aligned} X_{it} &= \boldsymbol{\lambda}_{it}' F_t + e_{it} \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{is})' F_t \\ &= \boldsymbol{\lambda}_{is}' F_t + e_{it} + w_{ist} \end{aligned}$$

- $w_{ist} \xrightarrow{p} 0$ as $(t-s)/T \to 0$
- For each $t \in \{s+1, \ldots, s+\tau\}$, if $\tau/T \to 0$,

$$X_{it} \approx \boldsymbol{\lambda}_{is}' F_t + e_{it}.$$

Estimating Factors

- Set $\tau = T/n$, where $n^{-1} + nT^{-1} \to 0$
- Split the data into n equal parts, $\mathcal{T}_r = \{(r-1)\tau + 1, \ldots, r\tau\}$
- For each $r = 1, \ldots, n$, compute

$$\min_{\boldsymbol{\Lambda}^{(r)}, \boldsymbol{F}^{(r)}} \frac{1}{N\tau} \sum_{i=1}^{N} \sum_{t \in \mathcal{T}_r} \left(X_{it} - \boldsymbol{\lambda}_i^{(r)\prime} \boldsymbol{F}_t^{(r)} \right)^2$$

Lemma

For each $t \in \mathcal{T}_r$, $r = 1, \ldots, n$, *and as* $N, T, n \to \infty$, *there exists a* $K \times K$ *matrix* $\boldsymbol{H}^{(r)}$ *with full rank such that*

$$\widetilde{\boldsymbol{F}}_t^{(r)} = \boldsymbol{H}^{(r)\prime} \boldsymbol{F}_t + O_p(N^{-1/2}) + O_p(\tau^{-1}) + O_p(n^{-\alpha})$$

Estimating Factors

Lemma

For each $t \in \mathcal{T}_r$, $r = 1, \ldots, n$, and as $N, T, n \to \infty$, there exists a $K \times K$ matrix $\mathbf{H}^{(r)}$ with full rank such that

$$\widetilde{\mathbf{F}}_t^{(r)} = \mathbf{H}^{(r)\prime} \mathbf{F}_t + \underbrace{O_p(N^{-1/2}) + O_p(\tau^{-1})}_{\text{estimation error}} + \; O_p(n^{-\alpha})$$

Estimating Factors

Lemma

For each $t \in \mathcal{T}_r$, $r = 1, \ldots, n$, and as $N, T, n \to \infty$, there exists a $K \times K$ matrix $\mathbf{H}^{(r)}$ with full rank such that

$$\widetilde{\mathbf{F}}_t^{(r)} = \mathbf{H}^{(r)\prime} \mathbf{F}_t + \underbrace{O_p(N^{-1/2}) + O_p(\tau^{-1})}_{\textit{estimation error}} + \underbrace{O_p(n^{-\alpha})}_{\substack{\textit{approximation} \\ \textit{error}}}$$

Identifying Factors

Following Bai and Ng (2013), we consider the following identification restrictions for each $r = 1, \ldots, n$ and for all t:

(PC1) $\tau^{-1} F^{(r)\prime} F^{(r)} = I$ and $\boldsymbol{\Lambda}_t' \boldsymbol{\Lambda}_t$ is a diagonal matrix with distinct entries, write $\widehat{F}^{(r)} = \widetilde{F}^{(r)}$;

(PC2) $T^{-1} F^{(r)\prime} F^{(r)} = I$ and $\boldsymbol{\Lambda}_t = \left(\boldsymbol{\Lambda}_{1t}', \boldsymbol{\Lambda}_{2t}'\right)'$ where $\boldsymbol{\Lambda}_{1t}$ is a lower triangular matrix with non-zero elements on the diagonal, write $\widehat{F}^{(r)} = \widetilde{F}^{(r)} Q^{(r)}$ where $\widetilde{\boldsymbol{\Lambda}}_r^{(r)\prime} = Q^{(r)} R^{(r)}$ is the QR decomposition;

(PC3) $\boldsymbol{\Lambda}_t = \left(I, \boldsymbol{\Lambda}_{2t}'\right)'$, while $F^{(r)}$ is not restricted, write $\widehat{F}^{(r)} = \widetilde{F}^{(r)} \widetilde{\boldsymbol{\Lambda}}_1^{(r)\prime}$.

Lemma

Under each of **PC1-PC3***, and for each t, we have*

$$\widehat{F}_t = F_t + O_p(C_{N\tau}^{-1}) + O_p(n^{-\alpha})$$

where $C_{N\tau} = \min\{\sqrt{N}, \sqrt{\tau}\}$.

Tasks

Questions to be asked (answered) in this paper:

1. ~~*Can we allow non-trivial variations in loadings and still estimate the factors without rotation?*~~
2. Can we estimate the sample paths of all loadings?
3. Can we estimate them more efficiently if some of the loadings are constant?
4. Can we identify them if some of the loadings are constant?

Estimating Factor Loadings

Approximate the time-varying loadings by a set of B-splines

$$\boldsymbol{\lambda}_{it} \approx C_i B_t$$

Estimating Factor Loadings

Approximate the time-varying loadings by a set of B-splines

$$\boldsymbol{\lambda}_{it} \approx \boldsymbol{C}_i \boldsymbol{B}_t$$

Then, we have

$$\begin{aligned} X_{it} &\approx (\boldsymbol{C}_i \boldsymbol{B}_t)' \boldsymbol{F}_t + e_{it} \\ &= \sum_{k=1}^{K} \sum_{j=-\kappa+1}^{n-1} c_{ikj} B_{jt} F_{kt} + e_{it} \\ &= \boldsymbol{c}_i' \mathbb{F}_t + e_{it} \end{aligned}$$

where $\mathbb{F}_t = \boldsymbol{F}_t \otimes \boldsymbol{B}_t$.

Estimating Factor Loadings

Approximate the time-varying loadings by a set of B-splines

$$\boldsymbol{\lambda}_{it} \approx \boldsymbol{C}_i \boldsymbol{B}_t$$

Then, we have

$$\begin{aligned} X_{it} &\approx (\boldsymbol{C}_i \boldsymbol{B}_t)' \widehat{\boldsymbol{F}}_t + e_{it} \\ &= \sum_{k=1}^{K} \sum_{j=-\kappa+1}^{n-1} c_{ikj} B_{jt} \widehat{F}_{kt} + e_{it} \\ &= \boldsymbol{c}_i' \widehat{\mathbb{F}}_t + e_{it} \end{aligned}$$

where $\widehat{\mathbb{F}}_t = \widehat{\boldsymbol{F}}_t \otimes \boldsymbol{B}_t$.

Estimating Factor Loadings

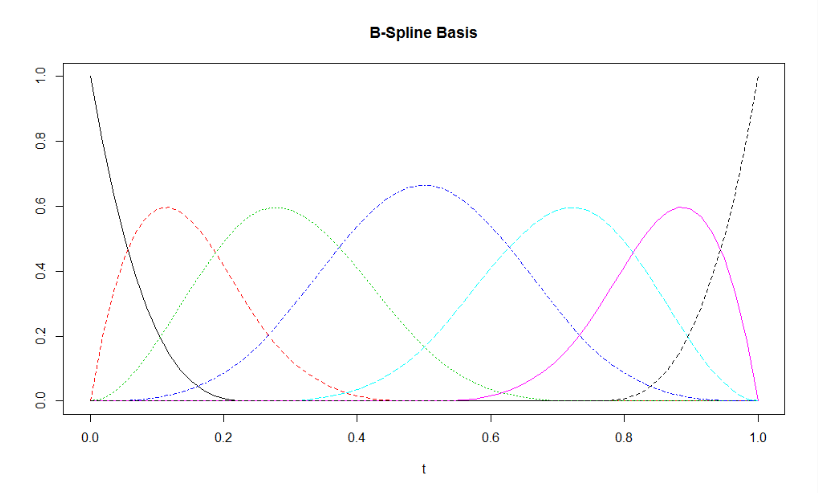
Consider the equidistant knot sequence $\boldsymbol{t} = (t_{-(d-1)}, \ldots, t_{n+\kappa-1})$

$$0 = t_{-(d-1)} = \cdots = t_0 < t_1 < \cdots < t_{n-1} < t_n = \cdots = t_{n+\kappa-1} = 1$$

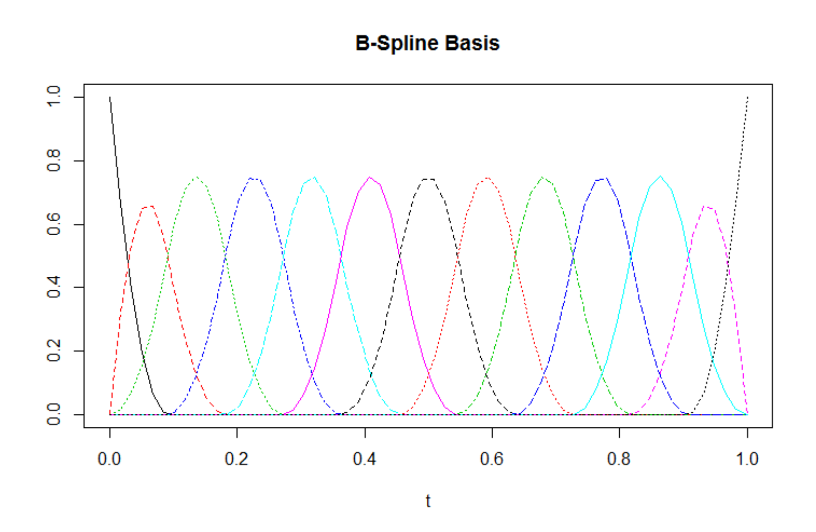
where $t_j = j/T$ for $j = 1, \ldots, n$. We focus on the B-splines of order κ defined equipped with $\boldsymbol{t}$, and define

$$\boldsymbol{B}(t) = (B_{-\kappa+1}, \ldots, B_{n-1})(t), \qquad \boldsymbol{B}_t = \boldsymbol{B}(t/T)$$

Estimating Factor Loadings



Estimating Factor Loadings



Estimating Factor Loadings

There exists a vector $\boldsymbol{c}_{ik}$ such that for each i, k,

$$\max_t |\lambda_{ikt} - \boldsymbol{c}_{ik}' \boldsymbol{B}_t| = O(n^{-\alpha})$$

Estimating Factor Loadings

There exists a vector $\boldsymbol{c}_{ik}$ such that for each i, k,

$$\max_t |\lambda_{ikt} - \boldsymbol{c}'_{ik}\boldsymbol{B}_t| = O(n^{-\alpha})$$

There exists matrix $\boldsymbol{C}_i$ and vector $\|\boldsymbol{\nu}_{it}\| = O(n^{-\alpha})$ such that

$$X_{it} = \boldsymbol{c}'_i\mathbb{F}_t + e_{it} + \boldsymbol{\nu}'_{it}\boldsymbol{F}_t$$

Estimating Factor Loadings

There exists a vector $\boldsymbol{c}_{ik}$ such that for each i, k,

$$\max_t |\lambda_{ikt} - \boldsymbol{c}_{ik}'\boldsymbol{B}_t| = O(n^{-\alpha})$$

There exists matrix $\boldsymbol{C}_i$ and vector $\|\boldsymbol{\nu}_{it}\| = O(n^{-\alpha})$ such that

$$X_{it} = \boldsymbol{c}_i'\mathbb{F}_t + e_{it} + \boldsymbol{\nu}_{it}'\boldsymbol{F}_t$$

Let $\widehat{\mathbb{F}}_t = \widehat{\boldsymbol{F}}_t \otimes \boldsymbol{B}_t$, we estimate

$$\widehat{\boldsymbol{c}}_i = \left(\widehat{\mathbb{F}}'\widehat{\mathbb{F}}\right)^{-1}\left(\widehat{\mathbb{F}}'X_i\right), \qquad \widehat{\boldsymbol{\lambda}}_{it} = \widehat{C}_i'\boldsymbol{B}_t$$

where $\widehat{\boldsymbol{c}}_i = \text{vec}\left(\widehat{\boldsymbol{C}}_i\right)$.

Estimating Factor Loadings

Theorem

For each i,

$$T^{-1} \sum_{t=1}^{T} \left\| \widehat{\boldsymbol{\lambda}}_{it} - \boldsymbol{\lambda}_{it} \right\|^2 = O_p(C_{N\tau}^{-2}) + O_p(n^{-2\alpha})$$

Moreover,

$$\sup_t \left\| \widehat{\boldsymbol{\lambda}}_{it} - \boldsymbol{\lambda}_{it} \right\| = O_p(\sqrt{n} C_{N\tau}^{-1}) + O_p(n^{1/2-\alpha})$$

Tasks

Questions to be asked (answered) in this paper:

1. ~~*Can we allow non-trivial variations in loadings and still estimate the factors without rotation?*~~
2. ~~*Can we estimate the sample paths of all loadings?*~~
3. Can we estimate them more efficiently if some of the loadings are constant?
4. Can we identify them if some of the loadings are constant?

Partially Linear Models

Consider the partially linear model,

$$X_{it} = \boldsymbol{\lambda}_{it}' F_{1t} + \boldsymbol{\gamma}_i' F_{2t} + e_{it}$$

Consider the partially linear model,

$$\begin{aligned} X_{it} &= \boldsymbol{\lambda}_{it}' \boldsymbol{F}_{1t} + \boldsymbol{\gamma}_i' \boldsymbol{F}_{2t} + e_{it} \\ &\approx \boldsymbol{c}_i' \widehat{\mathbb{F}}_{1t} + \boldsymbol{\gamma}_i' \widehat{\boldsymbol{F}}_{2t} + e_{it} \end{aligned}$$

Partially Linear Models

Consider the partially linear model,

$$\begin{aligned} X_{it} &= \boldsymbol{\lambda}_{it}' F_{1t} + \boldsymbol{\gamma}_i' F_{2t} + e_{it} \\ &\approx \boldsymbol{c}_i' \widehat{\mathbb{F}}_{1t} + \boldsymbol{\gamma}_i' \widehat{F}_{2t} + e_{it} \end{aligned}$$

We compute

$$\begin{pmatrix} \widehat{\boldsymbol{c}}_i \\ \widehat{\gamma}_i \end{pmatrix} = \left(\widehat{\mathbb{F}}' \widehat{\mathbb{F}} \right)^{-1} \left(\widehat{\mathbb{F}}' X_i \right)$$

where $\widehat{\mathbb{F}} = (\widehat{\mathbb{F}}_1, \widehat{F}_2)$.

Partially Linear Models

Theorem

For each i,

$$\|\widehat{\boldsymbol{\gamma}}_i - \boldsymbol{\gamma}_i\| = O_p(C_{N\tau}^{-1}) + O_p(n^{-\alpha})$$

Tasks

Questions to be asked (answered) in this paper:

1. ~~*Can we allow non-trivial variations in loadings and still estimate the factors without rotation?*~~
2. ~~*Can we estimate the sample paths of all loadings?*~~
3. ~~*Can we estimate them more efficiently if some of the loadings are constant?*~~
4. Can we identify them if some of the loadings are constant?

Determining Constancy of Factor Loadings

For each cross-section unit i and for each factor k, we choose between

Model 1 $\lambda_{iks} \neq \lambda_{ikt}$ for some $s \neq t$

Model 2 $\lambda_{ikt} = \gamma_{ik}$ for all t

Determining Constancy of Factor Loadings

For each cross-section unit i and for each factor k, we choose between

Model 1 $\lambda_{iks} \neq \lambda_{ikt}$ for some $s \neq t$

Model 2 $\lambda_{ikt} = \gamma_{ik}$ for all t

We compare

$$TVC_{1ik} = T^{-1} \sum_{t=1}^{T} (X_{it} - \widehat{\boldsymbol{\lambda}}_{it}' \widehat{\mathbf{F}}_t)^2 + g(T)$$

$$TVC_{2ik} = T^{-1} \sum_{t=1}^{T} (X_{it} - \widehat{\boldsymbol{\lambda}}_{-k,it}' \widehat{\mathbf{F}}_{-k,t} - \widehat{\gamma}_{ik} \widehat{F}_{kt})^2$$

Determining Constancy of Factor Loadings

Theorem

Let $g(T) \to 0$ *and* $(n^{\alpha} + C_{N\tau})g(T) \to \infty$*, then* $P(TVC_{1ik} < TVC_{2ik}) \to 1$ *under* ***Model 1*** *(time-varying loadings), and* $P(TVC_{1ik} > TVC_{2ik}) \to 1$ *under* ***Model 2*** *(constant loadings).*

SIMULATION RESULTS

Data Generating Processes

Denote $\mathfrak{B}_H(\cdot)$ as the fractional Brownian motion with Hurst index H,

$$
\begin{aligned}
X_{it} &= \lambda_{it} F_{1t} + \gamma_i F_{2t} + e_{it}, \qquad && e_{it} \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \\
F_{kt} &= \phi F_{k,t-1} + \varepsilon_{kt}, && \varepsilon_{kt} \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \\
\lambda_{it} &\sim \mathfrak{B}_{H_i}(t/T)
\end{aligned}
$$

Estimation Results: Latent Factors

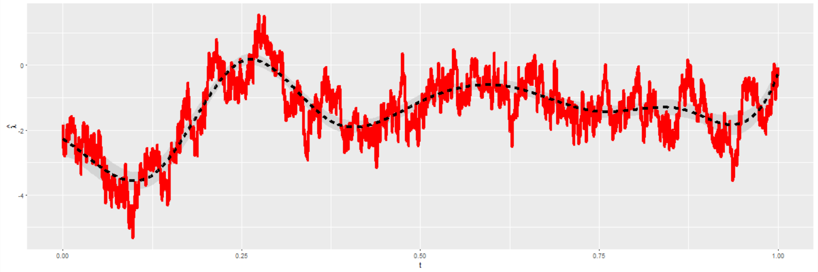
$N = 100$			$N = 500$		
T	H	$R^2_{\widehat{F}}$	T	H	$R^2_{\widehat{F}}$
400	0.3	0.956	400	0.3	0.957
400	0.5	0.965	400	0.5	0.972
400	0.8	0.966	400	0.8	0.972
1000	0.3	0.955	1000	0.3	0.982
1000	0.5	0.959	1000	0.5	0.984
1000	0.8	0.959	1000	0.8	0.984
2000	0.3	0.967	2000	0.3	0.988
2000	0.5	0.971	2000	0.5	0.989
2000	0.8	0.971	2000	0.8	0.988

Estimation Results: Factor Loadings

$N = 100$				$N = 500$			
T	H	$MSE(\lambda_{it})$	$MSE(\gamma_i)$	T	H	$MSE(\lambda_{it})$	$MSE(\gamma_i)$
400	0.3	0.688	0.008	400	0.3	0.690	0.011
400	0.5	0.239	0.005	400	0.5	0.231	0.005
400	0.8	0.099	0.005	400	0.8	0.096	0.004
1000	0.3	0.614	0.003	1000	0.3	0.611	0.003
1000	0.5	0.180	0.002	1000	0.5	0.175	0.002
1000	0.8	0.053	0.002	1000	0.8	0.049	0.002
2000	0.3	0.489	0.002	2000	0.3	0.514	0.001
2000	0.5	0.126	0.001	2000	0.5	0.131	0.001
2000	0.8	0.033	0.001	2000	0.8	0.034	0.001

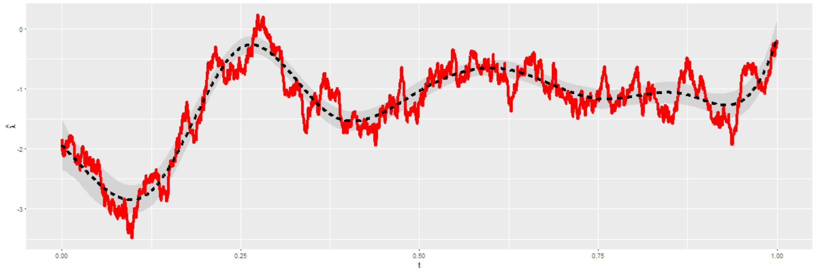
Estimation Results: Factor Loadings

$H = 0.3$



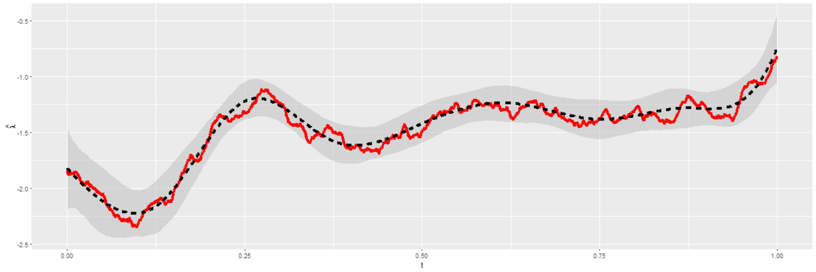
Estimation Results: Factor Loadings

$H = 0.5$



Estimation Results: Factor Loadings

$H = 0.8$



Determining Constancy of Loadings

$N = 100$				$N = 500$			
T	H	$P(M_1 \mid M_1)$	$P(M_2 \mid M_2)$	T	H	$P(M_1 \mid M_1)$	$P(M_2 \mid M_2)$
400	0.3	1.000	0.955	400	0.3	1.000	0.937
400	0.5	1.000	0.956	400	0.5	1.000	0.943
400	0.8	0.992	0.946	400	0.8	0.993	0.948
1000	0.3	1.000	0.970	1000	0.3	1.000	0.986
1000	0.5	1.000	0.981	1000	0.5	1.000	0.989
1000	0.8	1.000	0.984	1000	0.8	0.997	0.985
2000	0.3	1.000	0.996	2000	0.3	1.000	0.996
2000	0.5	1.000	0.998	2000	0.5	1.000	0.997
2000	0.8	0.999	0.998	2000	0.8	0.996	0.995

M_1 stands for varying loadings; M_2 stands for constant loadings.

Tasks

Questions to be asked (answered) in this paper:

1. Can we allow non-trivial variations in loadings and still estimate the factors without rotation?
2. Can we estimate the sample paths of all loadings?
3. Can we estimate them more efficiently if some of the loadings are constant?
4. Can we identify them if some of the loadings are constant?

Thanks!