Interpretable Sparse Proximate Factors for Large Dimensions

Markus Pelger [1] Ruoxuan Xiong [2]

Stanford University

9th Western Conference in Financial Mathematics
Nov 16, 2018

Motivation: What are the factors?

Statistical Factor Analysis

- Factor models are widely used in big data settings
  - Summarize information and reduce data dimensionality
  - Problem: Which factors should be used?
- Statistical (latent) factors perform well
  - Factors estimated from Principle Component Analysis (PCA)
  - Weighted averages of all cross-section units
  - Problem: Hard to interpret

Goals of this paper:

Create interpretable sparse proximate factors

- Shrink most small cross section units' weights to zero to get proximate factors

$\Rightarrow$ More interpretable

Contribution of this paper

Contribution

- This Paper: Estimation of interpretable proximate factors
- Key elements of estimator:
  1. Statistical factors instead of pre-specified (and potentially miss-specified) factors
  2. Uses information from large panel data sets: Many cross section units with many time observations
  3. Proximate factors approximate latent factors very well with a few cross section units without sparse structure in population loadings
  4. Only 5-10% of the cross-sectional observations with the largest exposure are needed for proximate factors

Contribution

Theoretical Results

- Asymptotic probabilistic lower bound for generalized correlations of proximate factors with population factors
- Guidance on how to construct proximate factors

Empirical Results

- Very good approximation to population factors with 5-10% cross-section units, measured by generalized correlation and variance explained
- Interpret statistical latent factors for
  - 370 single-sorted anomaly portfolios
  - 128 macroeconomic variables

Literature (partial list)

- Large-dimensional factor models with PCA
  - Bai and Ng (2002): Number of factors
  - Bai (2003): Distribution theory
  - Fan et al. (2013): Sparse matrices in factor modeling
  - Fan et al. (2016): Projected PCA for time-varying loadings
  - Pelger (2017), Aït-Sahalia and Xiu (2015): High-frequency
  - Kelly, Pruitt and Su (2017): IPCA
- Factor models with penalty term
  - Bai and Ng (2017): Robust PCA with ridge shrinkage
  - Lettau and Pelger (2018): Risk-Premium PCA with pricing penalty
  - Zhou et al. (2006): Sparse PCA (low dimension)

Single-sorted Portfolios

Portfolio Data

- Monthly return data from 07/1963 to 12/2016 ($T = 638$) for $N = 370$ portfolios
- Kozak, Nagel and Santosh (2017) data: 370 decile portfolios sorted according to 37 anomaly characteristics, such as momentum, volatility, turnover, size and volume.
- Estimate a 5-factor model similar as Lettau and Pelger (2018)

Single-sorted Portfolios: Fourth Factor

- Hard to interpret...

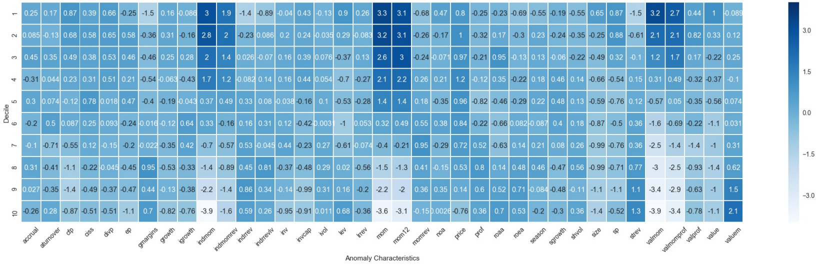


Figure: Financial single-sorted portfolios: Portfolio weights of 4th PCA factor.

Single-sorted Portfolios: Fourth Factor

- The fourth factor is a long-short momentum factor

⇒ Long-short extreme portfolios sorted by Industry Mom. Reversals, Momentum (6m), Momentum (12m), Value-Momentum, Value-Momentum-Prof. characteristics

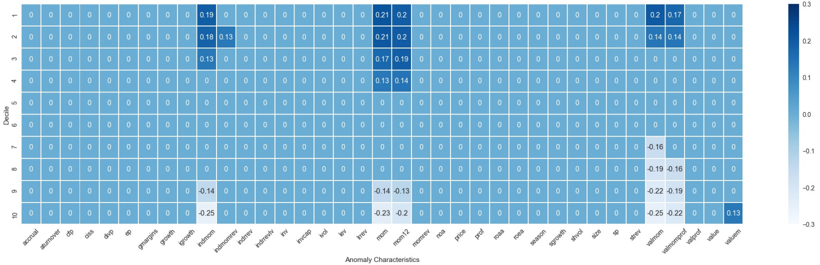


Figure: Financial single-sorted portfolios: Portfolio weights of 4th proximate factor. The sparse loading has 30 nonzero entries.

The Model

Approximate Factor Model

- Observe panel data of N cross-section units over T time periods:

$$X_{i,t} = \underbrace{\underset{1\times K}{\Lambda_i}^\top}_{loadings} \underbrace{\underset{K\times 1}{F_t}}_{factors} + \underbrace{e_{i,t}}_{idiosyncratic} \qquad i = 1, ..., N \quad t = 1, ..., T$$

- Matrix notation

$$\underbrace{X}_{N\times T} = \underbrace{\Lambda}_{N\times K} \underbrace{F^\top}_{K\times T} + \underbrace{e}_{N\times T}$$

  - N assets (large)
  - T time-series observation (large)
  - K systematic factors (fixed)
- F, Λ and e are unknown

The Model

Approximate Factor Model

- Systematic and non-systematic risk (F and e uncorrelated):

$$Var(X) = \underbrace{\Lambda Var(F)\Lambda^\top}_{systematic} + \underbrace{Var(e)}_{non-systematic}$$

$\Rightarrow$ Systematic factors should explain a large portion of the variance

$\Rightarrow$ Idiosyncratic risk can be weakly correlated

Estimation: PCA (Principal Component Analysis)

- Apply PCA to the sample covariance matrix: $\frac{1}{T}XX^\top - \bar{X}\bar{X}^\top$ with $\bar{X}$ = sample mean of asset excess returns
- Eigenvectors of largest eigenvalues estimate loadings $\hat{\Lambda}$
- $\hat{F}$ estimator for factors: $\hat{F} = X\hat{\Lambda}^\top(\hat{\Lambda}^\top\hat{\Lambda})^{-1} = \frac{1}{N}X^\top\hat{\Lambda}$

Method to Construct Sparse Proximate Factors

Proximate Factors

- Estimate loadings $\hat{\Lambda}$ by applying PCA to $\frac{1}{T}XX^\top - \bar{X}\bar{X}^\top$
- Sparse loadings $\tilde{\Lambda}$ are obtained from
  - Select finitely many m loadings with largest absolute value from $\hat{\Lambda}_k$ for all k
  - Shrink estimated loadings $\hat{\Lambda}$ to 0 except for m largest values
  - Divide by column norms, i.e. $\tilde{\lambda}_k^\top \tilde{\lambda}_k = 1$
- Proximate factors $\tilde{F} = X^T\tilde{\Lambda}(\tilde{\Lambda}^T\tilde{\Lambda})^{-1}$

Asymptotic results

- Proximate factors $\tilde{F}$ are in general not consistent.
  - Consider one-factor model

$$\tilde{F} = X^T \tilde{\Lambda} = F \Lambda^T \tilde{\Lambda} + e^T \tilde{\Lambda}$$

  - Idiosyncratic component not diversified away
  - Assume $e_{i,t} \overset{iid}{\sim} (0, \sigma^2_{e_{\cdot,t}})$, then each element in $e^T \tilde{\Lambda}$ has

$$Var\left(\sum_{i=1}^{m} \tilde{\lambda}_{1,1_i} e_{1_i,t}\right) = \sum_{i=1}^{m} \tilde{\lambda}^2_{1,1_i} \sigma^2_{e_{\cdot,t}} = \sigma^2_{e_{\cdot,l}} \not\to 0$$

Closeness between Proximate Factors and Latent Factors

Closeness measure

- For 1-factor model: Correlation between $\tilde{F}$ and F.
- Problem for multiple factors: Factors are only identified up to invertible linear transformations $\Rightarrow$ Need measure for closeness between span of two vector spaces
- For multi-factor model: The "closeness" between $\tilde{F}$ and F is measured by generalized correlation:
  - Total generalized correlation measure:
    $$\rho = trace\left((F^T F/T)^{-1}(F^T \tilde{F}/T)(\tilde{F}^T \tilde{F}/T)^{-1}(\tilde{F}^T F/T)\right)$$
    - $\rho = 0$: $\tilde{F}$ and F are orthogonal
    - $\rho = K$: $\tilde{F}$ and F span the same space

Intuition: Why does picking largest elements in $\hat{\Lambda}$ work?

- Consider one factor and one nonzero element in $\tilde{\Lambda}$: $F = [f_{1t}] \in \mathbb{R}^{T\times 1}$, $\Lambda = [\lambda_{1,i}] \in \mathbb{R}^{N\times 1}$
- $\tilde{\Lambda} = [\tilde{\lambda}_{1,i}]$ is sparse. Assume nonzero element in $\tilde{\lambda}_{1,i}$ is $\tilde{\lambda}_{1,1} = 1$, so $\tilde{\Lambda}^\top \tilde{\Lambda} = I$.

$$
\begin{aligned}
\tilde{F} &= X^T\tilde{\Lambda} = F\Lambda^T\tilde{\Lambda} + e^T\tilde{\Lambda} \\
&= f_1\lambda_{1,1} + e_1
\end{aligned}
$$

- Assume

$$
f_{1,t} \sim (0, \sigma_f^2), \qquad e_{1,t} \overset{iid}{\sim} (0, \sigma_e^2)
$$

$$
\frac{f_1^T f_1}{T} \to \sigma_f^2, \qquad \frac{e_1^T e_1}{T} \to \sigma_e^2
$$

- Define signal-to-noise ratio $s = \frac{\sigma_f}{\sigma_e}$

Intuition: Why pick the largest elements in $\hat{\Lambda}$?

$$\begin{aligned}
\rho &= tr\left((F^TF/T)^{-1}(F^T\tilde{F}/T)(\tilde{F}^T\tilde{F}/T)^{-1}(\tilde{F}^TF/T)\right) \\
&= \left(\frac{f_1^T(f_1\lambda_{1,1}+e_1)/T}{(f_1^Tf_1/T)^{1/2}((f_1\lambda_{1,1}+e_1)^T(f_1\lambda_{1,1}+e_1)/T)^{1/2}}\right)^2 \\
&\to \frac{\lambda_{1,1}^2}{\lambda_{1,1}^2+1/s^2}
\end{aligned}$$

- (Generalized) correlation increases in size of loading $|\lambda_{1,1}|$.
- (Generalized) correlation increases in signal-to-noise ratio s.
- No sparsity in population loadings assumed!

$\Rightarrow$ We provide probabilistic lower bound for (generalized) correlation ρ given a target correlation level ρ_0:

$$P(\rho > \rho_0)$$

Assumptions

Assumptions

Similar assumptions as in Bai and Ng (2002)

1. **Factors:** $E\left\|f_t\right\|^4 \leq M < \infty$ and $\frac{1}{T}\sum_{t=1}^{T} f_t f_t^T \xrightarrow{P} \Sigma_F$ for some $K \times K$ positive definite matrix $\Sigma_F = diag(\sigma_{f_1}^2, \sigma_{f_2}^2, \cdots, \sigma_{f_r}^2)$.
2. **Loadings:** Random variables $\max_i \|\lambda_{j,i}\| = O_p(1)$ and $\Lambda^\top \Lambda / N \to \Sigma_\Lambda$, independent of factors and errors
3. **Systematic factors:** Eigenvalues of $\Sigma_\Lambda \Sigma_F$ bounded away from 0 and ∞
4. **Residuals:** Weak Dependency
   - Bounded eigenvalues and sparsity of Σ_e
   - e weakly dependent with F
   - Light tails

$\Rightarrow$ Uniform convergence result for loadings $\forall i$, $\exists H$,

$$\max_{i \leq N} \left\| \hat{\lambda}_{(i)} - H\lambda_{(i)} \right\| = O_p\left(\frac{1}{\sqrt{N}} + \frac{N^{1/4}}{\sqrt{T}} \right).$$

One factor case: Extreme value theory

Theorem 1: Distribution of correlation

Assume: $K = 1$ factor and there exists sequences of constants $\{a_{1,N} > 0\}$ and $\{b_{1,N}\}$ such that

$$P((|\lambda_{1,(1)}| - b_{1,N})/a_{1,N} \leq z) \rightarrow G_1(z),$$

Then for $N, T \rightarrow \infty$

$$P(\rho \geq \rho_0) \geq 1 - G_{1,m}(z) + o_p(1)$$

$$\rho_0 = \frac{\sigma_{f_1}^2 (a_{1,N} z + b_{1,N})^2}{\frac{1+h(m)}{m}\sigma_e^2 + \sigma_{f_1}^2 (a_{1,N} z + b_{1,N})^2}$$

G_1 is the Generalized Extreme Value (GEV) distribution function,

$$G_1 = exp\left\{ -\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

One factor case: Extreme value theory

A few examples for G_1 and $a_{1,N}$ and $b_{1,N}$ for $\lambda_{1,i}$:

1. $G_1 \sim$ Gumbel distribution:
   - Standard normal distribution ($\lambda_i \sim N(0,1)$): $a_{1,N} = \frac{1}{N\phi(b_{1,N})}$ and $b_{1,N} = \Phi^{-1}(1-1/N)$, where $\phi(\cdot), \Phi(\cdot)$ are pdf and cdf of standard normal.
   - Exponential distribution ($\lambda_i \sim exp(1)$): $a_{1,N} = 1, b_{1,N} = N$
2. $G_1 \sim$ Frechet distribution:
   - $F_\lambda(x) = exp(-1/x)$: $a_{1,N} = N, b_{1,N} = 0$.
3. $G_1 \sim$ Weibull distribution:
   - Uniform: distribution ($\lambda_i \sim Uniform(0,1)$): $a_{1,N} = 1/N, b_{1,N} = 1$.

$\Rightarrow$ allows $\lambda_{1,i}$ to be cross-sectionally dependent, characterized by an extremal index θ, appeared in G_1

One factor case: Extreme value theory

For a target probability $p = 1 - G_{1,m}(z)$ (z can be calculated), the threshold $\rho_0 = \frac{\sigma_{f_1}^2 (a_{1,N}z + b_{1,N})^2}{\frac{1+h(m)}{m}\sigma_e^2 + \sigma_{f_1}^2 (a_{1,N}z + b_{1,N})^2}$ such that $P(\rho \geq \rho_0) \geq p + o_p(1)$ has

- ρ_0 increases in signal-to-noise ratio $s = \sigma_{f_1}/\sigma_e$
- ρ_0 increases in the dispersion of loadings' distribution
- ρ_0 increases in $\#$ nonzeros m and N (from simulation)
- ρ_0 decreases in $h(m)$ ($h(m)$ measures how correlated idiosyncratic errors are)

Multi Factors

Challenges

- Thresholded loadings/Proximate factors are in general not orthogonal to each other
- Generalized correlation take this into account

Additional Assumptions

1. Each cross section unit can only have very large exposure to one factor
2. Tail distributions for each factor loading are asymptotically independent

⇒ Only for theoretical derivation, not needed for this approach to work in simulation and empirical applications

⇒ Assumption 1 can be relaxed: some cross section units only have large exposure to one factor after rotated by some matrix

Multi Factors

Theorem 2: Distribution of generalized correlation

The asymptotic lower bound equals

$$\lim_{N,T\to\infty} P\left(\rho \geq \rho_0\right) \geq \prod_{j=1}^{K}\left(1 - G^*_{j,m}(\tau)\right) - \lim_{N\to\infty} P(\sigma_{\min}(B) < \underline{\gamma}) \qquad (1)$$

$$\rho_0 = K - \frac{(1+h(m))\sigma_e^2}{m\underline{\gamma}^2}\sum_{j=1}^{K}\frac{1}{s_j u^2_{j,N}(\tau)},$$

where $S = diag(s_1, s_2, \cdots, s_K)$ are the eigenvalues of $\Sigma_F\Sigma_\Lambda$ in decreasing order and $0 < \underline{\gamma} < 1$.

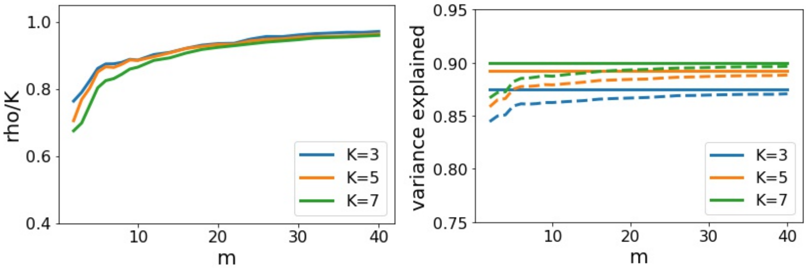
- ⇒ $\prod_{j=1}^{K}\left(1 - G^*_{j,m}(\tau)\right)$: product of loadings' tail distributions (asymptotically independent)
- ⇒ $B \propto S^{1/2}\Lambda^\top\tilde{\Lambda}$. $P(\sigma_{\min}(B) < \underline{\gamma})$: $\sigma_{\min}(B)$ measures how correlated one thresholded loading is to other population factor loadings

Single-sorted Portfolios

Portfolio Data

- Monthly return data from 07/1963 to 12/2016 ($T = 638$) for $N = 370$ portfolios
- Kozak, Nagel and Santosh (2017) data: 370 decile portfolios sorted according to 37 anomaly characteristics, such as momentum, volatility, turnover, size and volume.

Single-sorted Portfolios



(a) Generalized Correlation (b) Variance Explained

Figure: Financial single-sorted portfolios: Generalized correlation between $\tilde{F}$ and $\hat{F}$ normalized by K and proportion of variance explained by $\tilde{F}$ and $\hat{F}$ as a function of non-zero loading elements m , where K varies from 3 to 7. ($N = 370$, $T = 638$)

Single-sorted Portfolios

m	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_3$	$\hat{F}_4$	$\hat{F}_5$
10	0.993	0.992	0.771	0.918	0.837
20	0.995	0.948	0.883	0.949	0.890
30	0.996	0.965	0.935	0.966	0.910
40	0.997	0.971	0.958	0.975	0.923

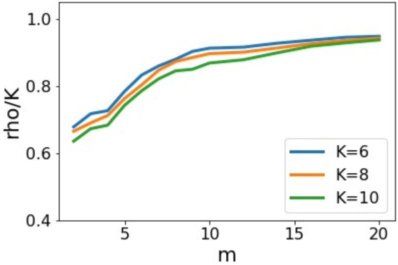
Table: Financial single-sorted portfolios: Generalized correlation between each $\hat{F}_j$ and all $\tilde{F}$ for $K = 5$. These generalized correlations correspond to R^2 from a regression of each $\hat{F}_j$ on all $\tilde{F}$.

Macroeconomic data

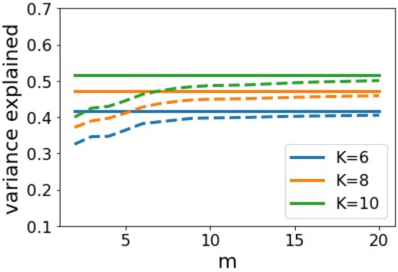
Macroeconomic Data

- Monthly data from 01/1959 to 02/2018 for $N = 128$ variables:
    1. output and income
    2. labor market
    3. housing
    4. consumption, orders and inventories
    5. money and credit
    6. interest and exchange rates
    7. prices
    8. stock market

Macroeconomic data



(a) Generalized Correlation



(b) Variance Explained

Figure: Macroeconomic data: Generalized correlation between $\tilde{F}$ and $\hat{F}$ normalized by K and proportion of variance explained by $\tilde{F}$ and $\hat{F}$, where K varies from 4 to 20. ($N = 128$, $T = 707$)

Macroeconomic data

m	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_3$	$\hat{F}_4$	$\hat{F}_5$	$\hat{F}_6$	$\hat{F}_7$	$\hat{F}_8$
10	0.953	0.959	0.949	0.953	0.961	0.799	0.833	0.767
15	0.967	0.970	0.958	0.956	0.964	0.857	0.867	0.837
20	0.977	0.974	0.957	0.963	0.961	0.905	0.919	0.891
25	0.983	0.980	0.961	0.979	0.973	0.937	0.943	0.929

Table: Macroeconomic data: Generalized correlation between each $\hat{F}_j$ and $\tilde{F}$, where $K = 8$. These generalized correlations correspond to R^2 from a regression of each $\hat{F}_j$ on all $\tilde{F}$.

Macroeconomic data

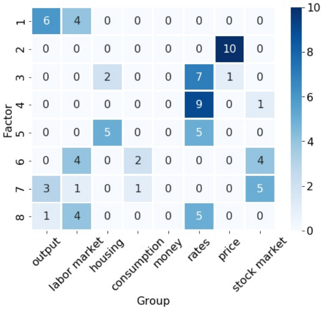


Figure: Macroeconomic data: 8-factor model, each sparse loading in $\tilde{\Lambda}$ has 10 nonzero entries. Values in this figure represent the number of nonzero entries in a particular group for a particular sparse loading. The 8 groups are: 1. output and income; 2. labor market; 3. housing; 4. consumption, orders and inventories; 5. money and credit; 6. interest and exchange rates; 7. prices; 8. stock market

Conclusion

Methodology

- Proximate factors (portfolios of a few cross-section units) for latent population factors (portfolios of all cross-section units)
- Simple thresholding estimator based on largest loadings
- Proximate factors approximate population factors well without sparsity assumption
- Asymptotic probabilistic lower bound for (generalized) correlation

$\Rightarrow$ A few observations summarize most of the information

Empirical Results

- Good approximation to population factors with 5-10% cross-section units

Multiple Factors

Multiple Factor: Rotate and threshold

- Assume there exists orthonormal matrix P s.t. large values in columns of $W^P = \Lambda HSP$ do not overlap (almost orthogonal)
- m nonzero entries in $\tilde{W}_j$ are the largest in $\hat{W}_j$ satisfying $\max_{j,k\neq j} |\hat{w}^P_{i,k}/\hat{w}^P_{i,j}| < c$ and are standardized by

$$\tilde{W}^P = \left[\frac{\hat{W}^P_1 \odot M_1}{\|\hat{W}^P_1 \odot M_1\|} \quad \frac{\hat{W}^P_2 \odot M_2}{\|\hat{W}^P_2 \odot M_2\|} \quad \cdots \quad \frac{\hat{W}^P_K \odot M_K}{\|\hat{W}^P_K \odot M_K\|} \right].$$

- The proximate factors are

$$\tilde{F}^P = X^T \tilde{W}^P ((\tilde{W}^P)^T \tilde{W}^P)^{-1} = X^T \tilde{W}^P$$

- Generalized Correlation

$$\rho = tr\left((F^T F/T)^{-1} (F^T \tilde{F}^P/T)((\tilde{F}^P)^T \tilde{F}^P/T)^{-1}((\tilde{F}^P)^T F/T) \right)$$

Multiple Factors

Theorem 4: Rotate and threshold

Let $\bar{w}^P_{(m),j}$ be the m-th order statistic of the entries in $|w^P_j|$ that satisfy $\max_{j,k\neq j} |w^P_{i,k}/w^P_{i,j}| < c$ and assume that the cumulative density function of $\bar{w}^P_{(m),j}$ is continuous. Then for a particular threshold $0 < \rho_0 < K$ and a fixed m, we have

$$\lim_{N,T\to\infty} P(\rho > \rho_0) \geq \lim_{N\to\infty} P\left(\sum_{j=1}^{K} \frac{1}{(\bar{w}^P_{(m),j})^2} < \frac{m(1-\gamma)(K-\rho_0)}{(1+f(m))\sigma_e^2}\right), \quad (2)$$

where $\gamma = c(2 + c(K-2))(K(K-1))^{1/2}$.

Relationship with Lasso

Alternative approach with Lasso:

1. Estimate factors by PCA, i.e $X^T X \hat{F} = \hat{F} V$ with V matrix of eigenvalues.
2. Estimate loadings by minimizing $\left\| X - \Lambda \hat{F}^T \right\|_F^2 + \alpha \left\| \Lambda \right\|_1$. Divide the minimizer by its column norm (standardize each loading) to obtain $\bar{\Lambda}$
3. Proximate factors from Lasso approach are $\bar{F} = X^T \bar{\Lambda} (\bar{\Lambda}^T \bar{\Lambda})^{-1}$

$\Rightarrow$ Same selection of non-zero elements (for one factor case) but different weighting

$\Rightarrow$ Under certain conditions worse performance than thresholding approach

- Tuning parameter less transparent

Simulation

- Compare probabilistic lower bounds with Monte-Carlo simulations
- **Factors:** $K = 1$ or $K = 2$ and $F_t \sim N(0, \sigma_f^2)$
- **Loadings:** $\lambda_i \sim N(0, 1)$ i.i.d.
- **Residuals:** $\sigma_e = 1$ and $e_{t,i} \sim N(0, 1)$ i.i.d.
- Vary signal-to-noise ratio with $\sigma_f \in \{0.8, 1.0, 1.2\}$
- $N = 100$) and $T \in \{50, 100, 200\}$
- We analyze:
  - Probabilistic lower bound for $\rho_0 = 0.95$
  - Distribution of lower bound with extreme value distribution

Simulation: One factor with very strong signal

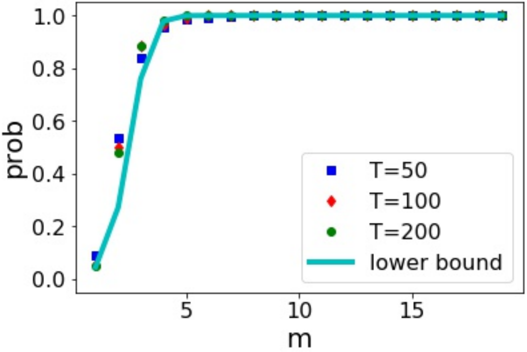


Figure: Probabilistic lower bound: $\sigma_f = 1.2$, $\rho_0 = 0.95$

Simulation: One factor with weaker signal

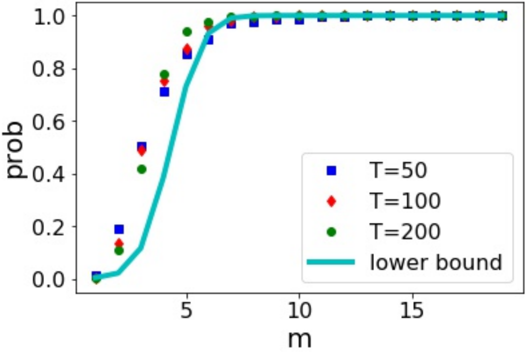


Figure: Probabilistic lower bound: $\sigma_f = 1.0$, $\rho_0 = 0.95$

Simulation: One factor with weak signal

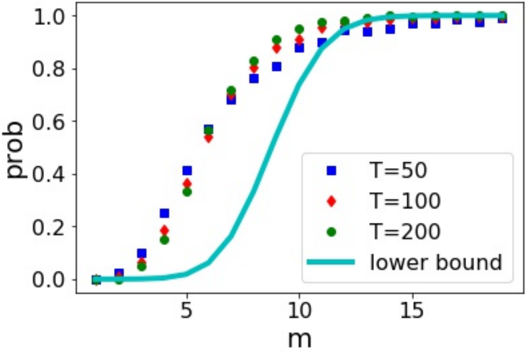
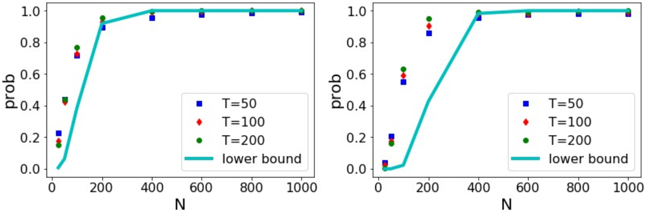


Figure: Probabilistic lower bound: $\sigma_f = 0.8$, $\rho_0 = 0.95$

Simulation: One factor with increasing N

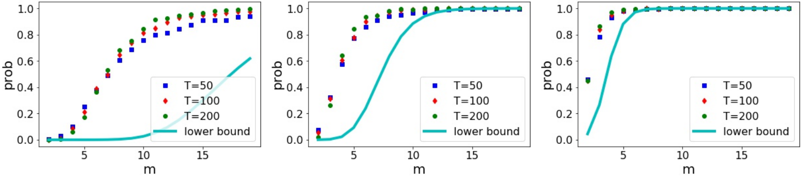


(a) One-factor model ($\sigma_f = 1.0$)

(b) Multi-factor model ($\sigma_f = [1.2, 1.0]$)

Figure: Probabilistic lower bound: $\rho_0 = 0.95$

Simulation: Two Factors

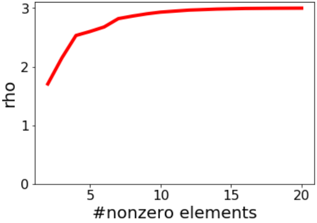


(a) $\sigma_f = [1.0, 0.8]$, (b) $\sigma_f = [1.2, 1.0]$ (c) $\sigma_f = [1.5, 1.2]$

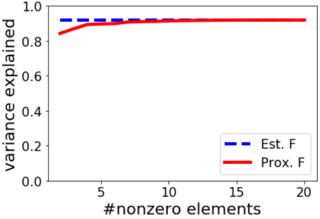
Figure: Probabilistic lower bound: $\rho_0 = 1.9$.

Empirical Application: Size and Investment Portfolios

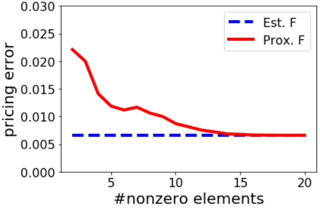
- 25 portfolios formed on size and investment (07/1963-10/2017, 3 factors, daily data)



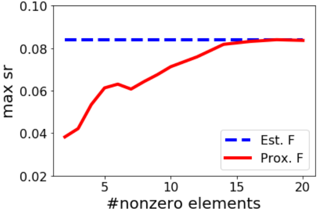
(a) Generalized correlation



(b) Variance explained



(c) RMS pricing error



(d) Max Sharpe Ratio

Empirical Application: Size and Investment Portfolios

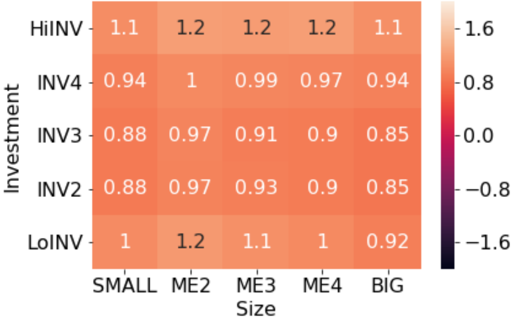


Figure: Portfolio weights of 1. statistical factor

$\Rightarrow$ Equally weighted market factor

Empirical Application: Size and Investment Portfolios

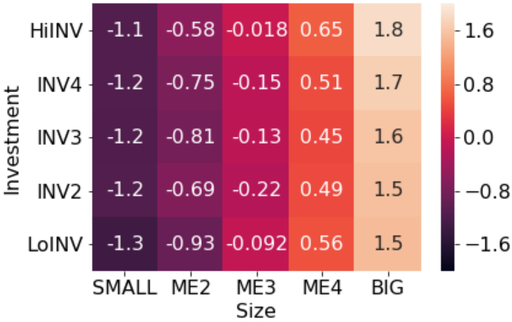


Figure: Portfolio weights of 2. statistical factor

⇒ Small-minus-big size factor

⇒ Proximate factor with 4 largest weights correlation 0.97 with size factor

Empirical Application: Size and Investment Portfolios

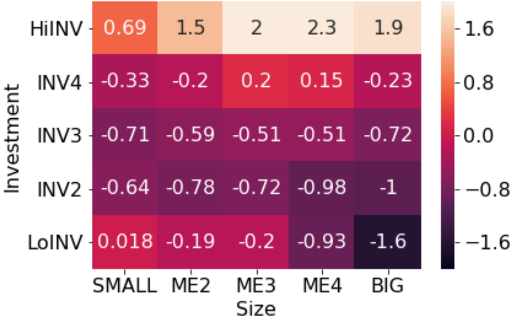


Figure: Portfolio weights of 3. statistical factor

⇒ High-minus-low value factor

⇒ Proximate factor with 4 largest weights correlation 0.79 with investment factor

Single-sorted portfolios

	Anomaly characteristics		Anomaly characteristics
1	Accruals - accrual	20	Momentum (12m) - mom12
2	Asset Turnover - aturnover	21	Momentum-Reversals - momrev
3	Cash Flows/Price - cfp	22	Net Operating Assets - noa
4	Composite Issuance - ciss	23	Price - price
5	Dividend/Price - divp	24	Gross Protability - prof
6	Earnings/Price - ep	25	Return on Assets (A) - roaa
7	Gross Margins - gmargins	26	Return on Book Equity (A) - roea
8	Asset Growth - growth	27	Seasonality - season
9	Investment Growth - igrowth	28	Sales Growth - sgrowth
10	Industry Momentum - indmom	29	Share Volume - shvol
11	Industry Mom. Reversals - indmomrev	30	Size - size
12	Industry Rel. Reversals - indrrev	31	Sales/Price - sp
13	Industry Rel. Rev. (L.V.) - indrrevlv	32	Short-Term Reversals - strev
14	Investment/Assets - inv	33	Value-Momentum - valmom
15	Investment/Capital - invcap	34	Value-Momentum-Prof. - valmomprof
16	Idiosyncratic Volatility - ivol	35	Value-Protability -valprof
17	Leverage - lev	36	Value (A) - value
18	Long Run Reversals - lrrev	37	Value (M) - valuem
19	Momentum (6m) - mom		