Arbitrage Portfolios

Soohun Kim Robert A. Korajczyk Andreas Neuhierl

December 2019

Georgia Tech

Northwestern University

University of Notre Dame

Outline

Motivation

- Large literature on cross-sectional return predictors (e.g. Cochrane (2011), Harvey (2015))
  - probably only few provide independent information for asset returns

Motivation

- Large literature on cross-sectional return predictors (e.g. Cochrane (2011), Harvey (2015))
  - probably only few provide independent information for asset returns
- Interpretation of predictors (characteristics) is often muddled or not discussed
  - proxy for factor loadings?
  - mispricing?
  - both?
  - often highly restricted tests in the literature

Motivation

- Large literature on cross-sectional return predictors (e.g. Cochrane (2011), Harvey (2015))
  - probably only few provide independent information for asset returns
- Interpretation of predictors (characteristics) is often muddled or not discussed
  - proxy for factor loadings?
  - mispricing?
  - both?
  - often highly restricted tests in the literature
- Most prevalent approach: portfolio constructions by double sorts on characteristic and associated factor beta (Daniel and Titman 1997)

Sorts of Sorts

Simulate CAPM return as $R_{i,t} = \beta_i f_{M,t} + e_{i,t}$

- Assume β_i is a linear function of characteristic, c_i
  - $\beta_i = \mu_\beta + \sigma_\beta c_i$ and $c_i \sim N(0,1)$
- Sort stock into portfolio based c_i and then on $\hat{\beta}_{i,t-1}$

			Past Beta										
			Low									High	
Characteristic			1	2	3	4	5	6	7	8	9	10	10-1
Low	1	0.24	0.25	0.29	0.34	0.26	0.15	0.34	0.25	0.18	0.14	0.23	-0.03
	2	0.36	0.40	0.32	0.39	0.36	0.37	0.39	0.29	0.32	0.25	0.47	0.06
	3	0.42	0.37	0.41	0.42	0.41	0.46	0.28	0.47	0.46	0.48	0.46	0.09
	4	0.45	0.46	0.45	0.45	0.36	0.42	0.44	0.39	0.45	0.58	0.47	0.02
	5	0.47	0.48	0.37	0.48	0.45	0.52	0.51	0.47	0.48	0.44	0.47	-0.01
	6	0.53	0.56	0.51	0.61	0.43	0.47	0.56	0.59	0.47	0.56	0.55	-0.01
	7	0.58	0.52	0.54	0.61	0.58	0.60	0.60	0.51	0.56	0.71	0.59	0.07
	8	0.59	0.56	0.54	0.56	0.63	0.60	0.46	0.66	0.70	0.62	0.56	0.01
	9	0.67	0.67	0.68	0.59	0.71	0.66	0.64	0.71	0.69	0.67	0.70	0.03
High	10	0.78	0.78	0.74	0.68	0.83	0.75	0.74	0.85	0.78	0.80	0.86	0.08
	10-1	0.54^{***}	0.52^{***}	0.45^{***}	0.33^{**}	0.57^{***}	0.61^{***}	0.40^{***}	0.60^{***}	0.60^{***}	0.66^{***}	0.63^{***}	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

- Large return spread between low and high characteristics
- Small return spread between low and high beta stocks
- Our method estimates $|\widehat{\theta}_\alpha| < 10^{-16}$ (p-value ≈ 0.82)

Goals

1. Propose a novel method
    - Individual stocks
    - Handle many characteristics
    - Identify relation between characteristics and both mispricing (α, if any) and risk (β) separately
    - Accommodate both priced and non-priced factors
    - Allow time-varying cross-sectional relation between characteristics and returns
        - Investors learn about predictors
        - Post publication anomalies often weaker (McLean and Pontiff (2016))

Goals

1. Propose a novel method
    - Individual stocks
    - Handle many characteristics
    - Identify relation between characteristics and both mispricing (α, if any) and risk (β) separately
    - Accommodate both priced and non-priced factors
    - Allow time-varying cross-sectional relation between characteristics and returns
        - Investors learn about predictors
        - Post publication anomalies often weaker (McLean and Pontiff (2016))
2. Provide simulation evidence in factor economies

Goals

1. Propose a novel method
    - Individual stocks
    - Handle many characteristics
    - Identify relation between characteristics and both mispricing (α, if any) and risk (β) separately
    - Accommodate both priced and non-priced factors
    - Allow time-varying cross-sectional relation between characteristics and returns
        - Investors learn about predictors
        - Post publication anomalies often weaker (McLean and Pontiff (2016))
2. Provide simulation evidence in factor economies
3. Apply estimator to a large panels of US stocks

Related Literature

1. Large panel factor estimation/pricing
    - Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988, 1993)
    - Stock and Watson (2002), Bai (2002), Zhang (2009), Gagliardini, Ossola, and Scaillet (2016), Kim and Skoulakis (2018), Kim, Korajczyk, Pukthuanthong and Roll (2019)

Related Literature

1. Large panel factor estimation/pricing
    - Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988, 1993)
    - Stock and Watson (2002), Bai (2002), Zhang (2009), Gagliardini, Ossola, and Scaillet (2016), Kim and Skoulakis (2018), Kim, Korajczyk, Pukthuanthong and Roll (2019)
2. Characteristics based (factor) models
    - Rosenberg (1974), Fama and French (1993), Daniel and Titman (1997)
    - Lewellen (2014), Green, Hand and Zhang (2017), Freyberger, Neuhierl and Weber (2019)
    - Connor, Hagmann and Linton (2012), Fan, Liao and Wang (2016), Kelly, Pruitt and Su (2017, 2018), Feng, Polson and Xu (2018), Chen, Liu, Wang, Wang and Yu (2019)

Preview of Results

- Our procedure performs well in disentangling alpha and beta in a simulated economy
- For U.S. Equities, characteristics are related to factor loadings
- For U.S. Equities, characteristics are also related to mispricing
- Arbitrage portfolios
  - Sharpe ratios > 1
  - Alphas $> 1\%$ per month against commonly used factor models
  - Returns decline only marginally over time
  - Small firms do not drive the results

Outline

Outline

Setup

Returns follow a K-factor model

$$R_{i,t} = \alpha_i + \boldsymbol{\beta}_i{}'\mathrm{f}_t + e_{i,t}$$
$$\mathrm{R} = \boldsymbol{\alpha} 1_T' + \mathrm{B}\mathrm{F}' + \mathrm{E}$$

- $\mathrm{R}_{N\times T}-$ excess returns, $\boldsymbol{\alpha}_{N\times 1}-$ mispricing term, $\mathrm{B}_{N\times K}-$factor loadings, $\mathrm{F}_{T\times K}-$factor returns
- $\boldsymbol{\alpha}$, B, F and E are not observed
  - Alphas the potential existence of mispricing
  - Many papers implicitly assume zero alpha in factor models
- Estimate mispricing component ($\boldsymbol{\alpha}$) and risk component (B)
- Study mispricing quantitatively
  - large asset market ($N \to \infty$)
  - short time-span (T fixed)

Setup

Characteristics useful to estimate factor loadings **and** possible mispricing

$$\boldsymbol{\alpha} = \mathrm{G}_{\boldsymbol{\alpha}}(\mathrm{X}) + \Gamma_{\alpha}$$
$$\mathrm{B} = \mathrm{G}_{\boldsymbol{\beta}}(\mathrm{X}) + \Gamma_{\beta},$$

- *Factor loading function* ($\mathrm{G}_{\beta}(\mathrm{X})$) maps characteristics into **betas**
  - risk related component of characteristics
- *Mispricing function* ($\mathrm{G}_{\alpha}(\mathrm{X})$) maps characteristics into **alphas**
  - mispricing related component of characteristics

Main Assumptions

Existence

$$\frac{\mathsf{G}_\alpha(\mathsf{X})'\mathsf{G}_\alpha(\mathsf{X})}{N} \to \delta \geq 0 \text{ as } N \to \infty$$

- Possible deviation from APT (finite squared pricing errors)
- APT implies $\frac{1}{N}\boldsymbol{\alpha}'\boldsymbol{\alpha} \to 0$ and, therefore, $\frac{\mathsf{G}_\alpha(\mathsf{X})'\mathsf{G}_\alpha(\mathsf{X})}{N} \to 0$ as $N \to \infty$
- Non-zero alpha is an **empirical question**

Main Assumptions

Existence

$$\frac{\mathsf{G}_{\alpha}(\mathsf{X})'\mathsf{G}_{\alpha}(\mathsf{X})}{N} \to \delta \geq 0 \text{ as } N \to \infty$$

- Possible deviation from APT (finite squared pricing errors)
- APT implies $\frac{1}{N}\boldsymbol{\alpha}'\boldsymbol{\alpha} \to 0$ and, therefore, $\frac{\mathsf{G}_{\alpha}(\mathsf{X})'\mathsf{G}_{\alpha}(\mathsf{X})}{N} \to 0$ as $N \to \infty$
- Non-zero alpha is an **empirical question**

Identification

$$\frac{\mathsf{G}_{\beta}(\mathsf{X})'\mathsf{G}_{\alpha}(\mathsf{X})}{N} \to 0_K \text{ as } N \to \infty.$$

- Separate risk from mispricing
- Maximum explanatory power to the factors

Outline

Constructing the Arbitrage Portfolio Extension of Fan, Liao, Wang (2016)

1. Estimate the factor loading function, $G_\beta(X)$, from the *de-meaned & projected* returns
    - Alpha drops out
    - Use information in characteristics to estimate loadings

2. Estimate the mispricing function, $G_\alpha(X)$, from the *average* returns
    - Combination of characteristics to explain average returns **beyond risk**

3. Construct portfolio weights and hold arbitrage portfolio *out-of-sample*
    - Empirically, is arbitrage meaningful, quantitatively?

Step 1 - Estimating the Factor Loading Function $G_\beta(X)$

De-mean returns $(J_T = I_T - \frac{1}{T} 1_T 1_T)'$

$$\begin{aligned} RJ_T &= \boldsymbol{\alpha} 1_T' J_T + BF'J_T + EJ_T \\ &= BF'J_T + EJ_T \\ &= (G_\beta(X) + \Gamma_\beta) F'J_T + EJ_T, \end{aligned}$$

- Alpha drops out (✓)
- De-meaned returns only related to risk (✓)
- Short time series (✗)
- Many loadings to estimate (✗)

Step 1 - Estimating the Factor Loading Function $G_\beta(X)$

Use extra information in characteristics

- Characteristics and loadings are related
- Project de-meaned returns onto characteristics ($P = X(X'X)^{-1}X'$)

$$\begin{aligned}\widehat{R} \equiv P(RJ_T) &= P\left(G_{\boldsymbol{\beta}}(X)F'J_T + \Gamma_\beta F'J_T + EJ_T\right)\\ &\approx G_{\boldsymbol{\beta}}(X)F'J_T.\end{aligned}$$

Consistency for $G_\beta(X)$

- Apply principal components (PCA) to $\frac{\widehat{R}\widehat{R}'}{N}$
- $\widehat{G}_\beta(X) \xrightarrow{p} G_\beta(X)$ as $N \to \infty$ (in MSE)
- k-th column of $\widehat{G}_\beta(X)$ – the k-th largest eigenvector of $\frac{\widehat{R}\widehat{R}'}{N}$

Step 2 - Estimating the Mispricing Function $G_\alpha(X)$

Average returns

$$
\begin{aligned}
\overline{R} = \frac{1}{T}R1_T &= \boldsymbol{\alpha}\frac{1}{T}1_T'1_T + B\frac{1}{T}F'1_T + \frac{1}{T}E'1_T \\
&= \left(G_{\boldsymbol{\alpha}}(X) + \Gamma_\alpha\right)\frac{1}{T}1_T'1_T + \left(G_{\boldsymbol{\beta}}(X) + \Gamma_\beta\right)\frac{1}{T}F'1_T + \frac{1}{T}E'1_T \\
&= G_{\boldsymbol{\alpha}}(X) + \Gamma_\alpha + \left(G_{\boldsymbol{\beta}}(X) + \Gamma_\beta\right)\overline{F} + \overline{E}.
\end{aligned}
$$

- Average returns related to risk factors **and** potential mispricing (✗)
- Cannot follow the same strategy as before (✗)
- Use information in previously estimated factor loadings

Step 2 - Estimating the Mispricing Function $G_\alpha(X)$.

Use constrained least squares to estimate $G_\alpha(X)$

- Cross-sectional variation in return explained by characteristics
    - NOT explained by risk
    - Orthogonal to estimated factor loadings

Consistency for $G_\alpha(X)$

Solve

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left(\overline{R} - X\boldsymbol{\theta}\right)' \left(\overline{R} - X\boldsymbol{\theta}\right) \quad \textit{subject to} \quad \widehat{G}_\beta(X)' X\boldsymbol{\theta} = 0_K,$$

- $\widehat{G}_\alpha(X) = X\widehat{\boldsymbol{\theta}}$
- $\widehat{G}_\beta(X)$ *is from before*
- $\widehat{G}_\alpha(X) \xrightarrow{p} G_\alpha(X)$ *as* $N \to \infty$ *(in MSE)*
- Equivalent Construction

Step 3 - The Arbitrage Portfolio

Would like to have the infeasible portfolio $w = \frac{1}{N} G_\alpha(X)$, because

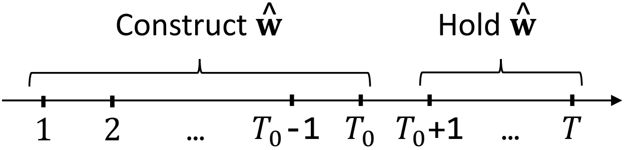
$$
\begin{aligned}
w'R = & \left(\overbrace{\frac{1}{N} G_\alpha(X)' G_\alpha(X)}^{\to \delta} + \overbrace{\frac{1}{N} G_\alpha(X)' \Gamma_\alpha}^{\to 0} \right) 1'_T \\
& + \left(\underbrace{\frac{1}{N} G_\alpha(X)' G_\beta(X)}_{\to 0} + \underbrace{\frac{1}{N} G_\alpha(X)' \Gamma_\beta}_{\to 0} \right) F' + \underbrace{\frac{1}{N} G_\alpha(X)' E}_{\to 0}.
\end{aligned}
$$

Consistency for Arbitrage Profits

- $\widehat{w} = \frac{1}{N} \widehat{G}_\alpha(X)$
- $\widehat{w}'R \xrightarrow{p} \delta 1'_T$

Outline

Overview of our out-of-sample strategy



1. Theory does not require large T
2. Estimate w over one sample ($t = 1, \cdots, T_0$) and calculate out-of-sample returns over a subsequent sample ($t = T_0 + 1, \cdots, T$)
3. We use $T_0 = 12$ and $T = 13$ for a baseline result

Calibration

1. Calibration period: 2011-2013
2. For the matrix X, we consider 61 characteristics (as in Freyberger et. al. (2019)), which are available at the beginning of 2011
3. We take a stance on the return generating process by considering four popular asset pricing models of CAPM, FF3, HXZ4, and FF5
   3.1 We use 2,458 individual stocks with full time series over the calibration sample period
   3.2 We calibrate α_i, $\boldsymbol{\beta}_i$, and the variance of residual returns, $\sigma^2_{i,\varepsilon} = \mathbb{E}\left[\varepsilon^2_{i,t}\right]$, of individual stocks
   3.3 After estimating $\widehat{\alpha}_i$ from time series regression, we fit the cross-sectional relation $\widehat{\alpha}_i = \mathrm{x}_i\widehat{\boldsymbol{\theta}} + e_i$ and rescale $\widetilde{\alpha}_i = k\widehat{\alpha}_i$ where $k = \frac{0.01}{\sqrt{\frac{\widehat{\boldsymbol{\theta}}'\mathrm{X}'\mathrm{X}\widehat{\boldsymbol{\theta}}}{N}}}$ so that $\frac{\mathrm{G}_\alpha(\mathrm{X})'\mathrm{G}_\alpha(\mathrm{X})}{N} \rightarrow$1 b.p./month

Simulation

1. In each repetition, we simulate returns from

$$\mathrm{R} = \boldsymbol{\alpha} 1_T' \sqrt{\delta} + \mathrm{BF}' + \mathrm{E},$$

   where $\boldsymbol{\alpha}$ and B are calibrated to match the properties of the CRSP data, F are bootstrapped from the realized factors over the 600-month sample from January 1967 to December 2016, E are drawn from a normal distribution with the calibrated parameters as on the previous slide
2. We consider $\delta = 0,\ 5,$ and 10
3. $\widehat{\mathrm{w}}$ is estimated with the returns over $t = 1, \cdots, 12$ and the return of the arbitrage portfolio is measured in the following month over $t = 13$
4. 10,000 repetitions

Simulations

1. True Model Known - ▸ base case
2. Too Many Factors - ▸ too many factors
3. Too Few Factors - ▸ too few factors
4. Time-Varying Characteristics – ▸ time varying characteristics
5. Correlated errors (industry cluster) – ▸ Correlated Errors
6. Different Calibration Period (financial crisis) – ▸ Different Calibration Period
7. Omitted Characteristics – ▸ Omitted Characteristics

Outline

Data

- CRSP and Compustat to create 61 characteristics (Freyberger et. al. (2019))
- Sample period: 1965 to 2018
- Characteristics: market cap, book-to-market, profitability, investment, beta, idiosyncratic volatility, turnover, bid-ask spread, short-term reversal, momentum, intermediate momentum, long-run reversal, total assets, cash over assets, D&A over assets, fixed costs to assets, capex to assets, operating leverage, price-to-cost margin, return-on-equity, operating accruals, free-cash flow to book value of equity, Tobin's Q, net payout ratio, assets-to-market cap, total assets, capital turnover, capital intensity, change in PP&E, earnings to price, and others

Empirical Implementation

1. Parametric assumptions
   - $G_\alpha(X)$ and $G_\beta(X)$ linear
   - Nonlinear expansions in robustness
2. Choices
   - 12-month moving estimation window
   - 1 month holding period
   - number of estimated factors (# eigenvectors) for $G_\beta(X)$ estimation
   - scale the estimation period volatility of $\hat{w}'R$ to 20% annualized

Annualized Performance Statistics

# Eigenvectors	Mean (%)	Standard Deviation (%)	Sharpe Ratio	Skewness	Kurtosis	Maximum Drawdown	Worst Month (%)	Best Month (%)
1	20.08	14.70	1.37	1.24	8.30	22.37	-18.77	29.46
2	24.84	17.51	1.42	0.53	6.07	23.16	-22.44	30.06
3	24.02	14.66	1.64	1.28	10.23	20.91	-19.89	34.05
4	27.54	16.56	1.66	1.07	6.71	22.21	-19.61	30.66
5	28.71	17.94	1.60	1.10	7.11	20.08	-20.08	36.16
6	29.48	18.42	1.60	1.29	8.67	20.84	-19.99	41.88
7	30.13	18.31	1.65	1.34	9.04	21.92	-20.21	42.82
8	29.67	19.84	1.50	1.26	10.38	27.92	-26.02	42.74
9	28.75	17.74	1.62	1.32	8.66	27.05	-20.43	36.56
10	24.93	19.02	1.31	0.29	15.72	38.52	-38.52	41.19

- portfolio properties change from one through five eigenvectors
- relatively similar portfolios for more than five eigenvectors
- maximum drawdown of common factors
    - 55.68% (market), 55.04% (size), 40.92% (value) and 57.31% (momentum)

Risk Adjusted Returns - 1 Eigenvector

	CAPM	FF3	FF3+UMD	FF5	FF5+UMD	HXZ4	HXZ4+UMD
alpha	1.64^{***}	1.57^{***}	1.38^{***}	1.70^{***}	1.53^{***}	1.58^{***}	1.54^{***}
	(0.22)	(0.21)	(0.20)	(0.23)	(0.23)	(0.25)	(0.25)
mktrf	0.06	0.01	0.06	0.00	0.03		
	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)		
smb		0.36^{**}	0.36^{**}	0.22	0.22^{*}		
		(0.17)	(0.14)	(0.14)	(0.11)		
hml		0.13	0.22	0.05	0.18		
		(0.13)	(0.14)	(0.14)	(0.12)		
umd			0.23^{*}		0.25^{**}		0.33^{***}
			(0.12)		(0.11)		(0.11)
rmw				-0.46^{***}	-0.51^{***}		
				(0.17)	(0.14)		
cma				0.17	0.06		
				(0.21)	(0.19)		
mkt						0.03	0.05
						(0.07)	(0.06)
me						0.29^{*}	0.21^{*}
						(0.17)	(0.11)
ia						0.21	0.22
						(0.22)	(0.20)
roe						-0.15	-0.44^{***}
						(0.16)	(0.14)
Adj. R^2	0.00	0.06	0.11	0.11	0.17	0.06	0.14
Num. obs.	612	612	612	612	612	612	612

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Risk Adjusted Returns - 6 Eigenvectors

	CAPM	FF3	FF3+UMD	FF5	FF5+UMD	HXZ4	HXZ4+UMD
alpha	2.49***	2.43***	2.00***	2.52***	2.17***	2.26***	2.18***
	(0.28)	(0.27)	(0.23)	(0.30)	(0.26)	(0.32)	(0.28)
mktrf	−0.06	−0.10	−0.01	−0.10	−0.03		
	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)		
smb		0.31	0.32**	0.17	0.15		
		(0.20)	(0.15)	(0.15)	(0.12)		
hml		0.12	0.30*	−0.04	0.25*		
		(0.13)	(0.15)	(0.17)	(0.13)		
umd			0.50***		0.52***		0.61***
			(0.12)		(0.10)		(0.11)
rmw				−0.47**	−0.59***		
				(0.20)	(0.15)		
cma				0.32	0.10		
				(0.24)	(0.22)		
mkt						−0.06	−0.01
						(0.07)	(0.06)
me						0.31	0.17
						(0.20)	(0.13)
ia						0.34	0.36
						(0.26)	(0.24)
roe						0.02	−0.53***
						(0.18)	(0.15)
Adj. R^2	0.00	0.03	0.18	0.07	0.23	0.04	0.21
Num. obs.	612	612	612	612	612	612	612

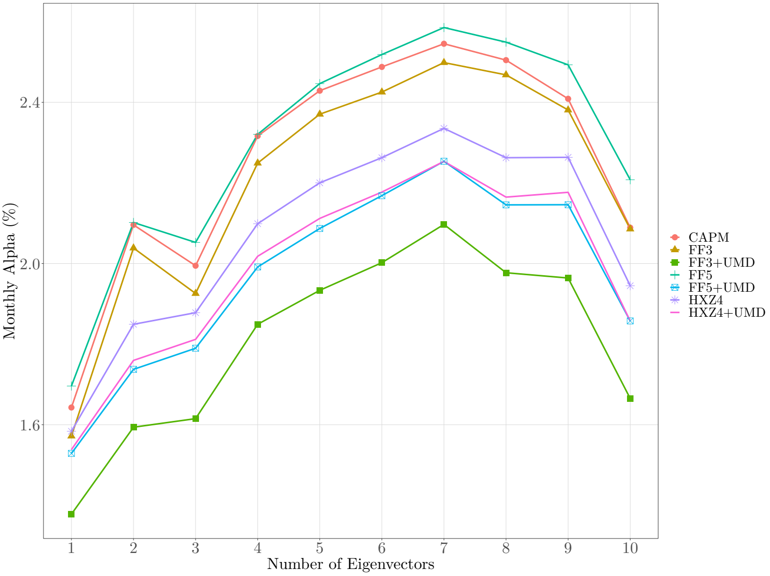
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Many, many Factors

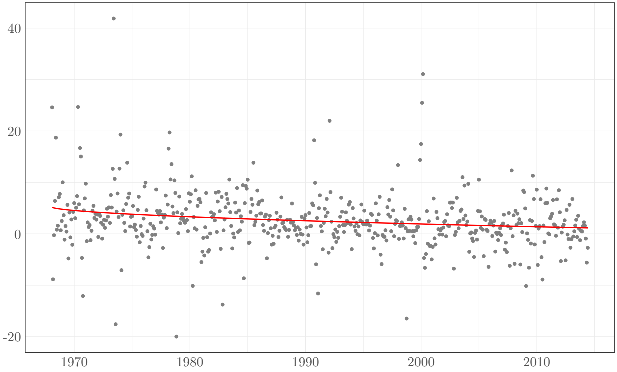
	CAPM	FF3	FF3+UMD	FF5	FF5+UMD	HXZ4	HXZ4+UMD
alpha	1.48^{***}	1.42^{***}	1.39^{***}	1.59^{***}	1.55^{***}	1.58^{***}	1.56^{***}
	(0.26)	(0.23)	(0.24)	(0.24)	(0.25)	(0.26)	(0.27)
mktrf	0.26^{***}	0.12	0.13^{*}	0.08	0.09		
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)		
smb		0.75^{***}	0.75^{***}	0.65^{***}	0.65^{***}		
		(0.17)	(0.16)	(0.18)	(0.18)		
hml		0.06	0.07	0.13	0.16		
		(0.13)	(0.12)	(0.16)	(0.15)		
umd			0.04		0.06		0.14
			(0.11)		(0.10)		(0.11)
rmw				-0.37^{***}	-0.38^{***}		
				(0.13)	(0.13)		
cma				-0.17	-0.19		
				(0.20)	(0.21)		
mkt						0.11	0.12
						(0.07)	(0.07)
me						0.64^{***}	0.60^{***}
						(0.18)	(0.18)
ia						-0.09	-0.09
						(0.17)	(0.17)
roe						-0.28^{*}	-0.41^{***}
						(0.15)	(0.15)
Adj. R^2	0.03	0.15	0.15	0.16	0.16	0.15	0.15
Num. obs.	612	612	612	612	612	612	612

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Alpha for Alternative Numbers of Factors

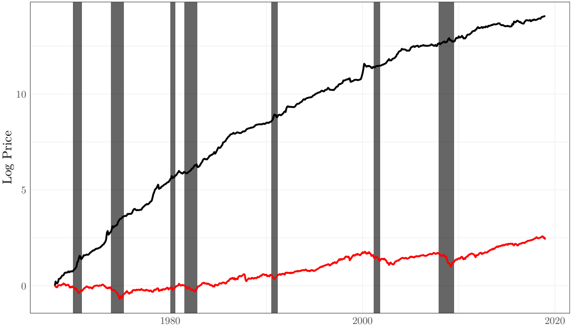


G_α (X) - **Declining Returns?**



- Test for time trend: $r_t = a + b \times t^\gamma + \varepsilon_t$.
- Point estimates
  - $\hat{a} = 5.22\%$ (p-value<0.01).
  - $\hat{b} = -0.11$ (insignificant) and $\hat{\gamma} = .57$ (insignificant)

Properties of the Arbitrage Portfolio - Price Path



- Black line - arbitrage portfolio, Red line - market return
- Arbitrage portfolio not strongly related to economy
- Relation to NBER recessions: $r_t = \alpha + \beta \times \text{NBER}_t + \varepsilon_t$
  - $\hat{\alpha} = 2.36\%$ (p-value < 0.01)
  - $\hat{\beta} = 0.71\%$ (p-value $= 0.26$)

Higher Bar - Statistical Factors

- Asymptotic Principal Components (Connor, Korajczyk (1986))
- Risk Premium Principal Components (Lettau, Pelger (2019))
- Instrumented Principal Components (Kelly, Pruitt, Su (2019))

	APC				RP-PC				IPCA			
	1 PC	5 PCs	10 PCs	20 PCs	1 RP-PC	5 RP-PCs	10 RP-PCs	20 RP-PCs	1 KPS-PC	5 KPS-PCs	10 KPS-PCs	20 KPS-PCs
	Panel A: Arbitrage Portfolio											
alpha	2.45***	2.08***	1.28***	1.12***	2.45***	1.15***	1.06***	1.01***	2.39***	2.31***	2.15***	1.57***
	(0.29)	(0.25)	(0.26)	(0.25)	(0.29)	(0.30)	(0.27)	(0.26)	(0.25)	(0.26)	(0.42)	(0.33)
Adj. R^2	-0.00	0.29	0.34	0.37	-0.00	0.31	0.34	0.36	0.02	0.13	0.14	0.23
Num. obs.	612	612	612	612	612	612	612	612	612	612	612	612

- Arbitrage Portfolio passes higher bar
- Economically and statistically significant alphas

Time Variation (FF5+UMD-α)

Lag in X	**Lag in θ_α**											
	1	2	3	4	5	6	7	8	9	10	11	12
1	2.17	1.45	1.26	1.32	1.35	1.31	1.39	1.03	1.25	1.38	1.30	1.27
2	1.42	1.00	0.73	0.81	0.52	0.60	0.61	0.42	0.49	0.64	0.39	0.49
3	1.08	0.70	0.64	0.66	0.71	0.56	0.55	0.31	0.31	0.38	0.25	0.24
4	1.14	0.80	0.67	0.85	0.67	0.70	0.68	0.42	0.47	0.58	0.50	0.57
5	1.14	0.80	0.68	0.83	0.75	0.73	0.74	0.47	0.49	0.53	0.49	0.50
6	1.14	0.75	0.68	0.72	0.73	0.84	0.80	0.52	0.47	0.50	0.43	0.52
7	1.11	0.76	0.82	0.79	0.80	0.89	0.97	0.62	0.62	0.62	0.45	0.46
8	1.17	0.76	0.75	0.66	0.73	0.76	0.88	0.59	0.56	0.56	0.40	0.47
9	1.00	0.62	0.66	0.54	0.57	0.60	0.73	0.42	0.39	0.42	0.26	0.25
10	1.11	0.72	0.59	0.71	0.61	0.73	0.81	0.54	0.54	0.66	0.45	0.52
11	1.17	0.83	0.73	0.81	0.75	0.80	0.85	0.62	0.66	0.64	0.63	0.63
12	1.23	0.78	0.77	0.81	0.60	0.69	0.77	0.49	0.40	0.53	0.41	0.58

- Each period we can update
  - $\hat{\theta}_\alpha$ – the cross-sectional relationship between characteristics and α (and β)
  - X – the characteristics (time-varying betas)

Influence of Small Firms (I)

- Drop all firms below the 10% NYSE market cap quantile
- Reduces sample size from 1.75 million to about 1 million

# Eigenvectors	Mean (%)	Standard Deviation (%)	Sharpe Ratio	Skewness	Kurtosis	Maximum Drawdown	Worst Month (%)	Best Month (%)
1	15.76	15.19	1.04	0.64	7.18	22.15	-21.62	31.03
2	19.92	18.42	1.08	0.27	5.18	25.33	-25.33	31.05
3	19.25	14.91	1.29	0.82	7.07	22.52	-22.52	31.49
4	22.00	17.45	1.26	0.68	5.49	23.80	-21.63	34.55
5	22.52	18.70	1.20	0.53	5.51	30.49	-21.93	36.71
6	22.01	19.31	1.14	0.38	5.49	27.17	-25.96	36.41
7	23.64	20.15	1.17	0.76	6.47	27.56	-27.56	41.44
8	23.25	20.61	1.13	0.79	6.75	28.05	-28.05	43.37
9	19.26	20.20	0.95	-0.51	12.34	48.41	-37.55	41.01
10	16.69	20.17	0.83	0.33	8.52	58.59	-28.68	42.82

Influence of Small Firms (II)

- Only slightly slower alphas
- Results not driven materially by small firms

# Eigenvectors	CAPM	FF3	FF3+UMD	FF5	FF5+UMD	HXZ4	HXZ4+UMD
1	1.27	1.21	0.95	1.34	1.12	1.16	1.10
2	1.65	1.62	1.09	1.74	1.30	1.38	1.28
3	1.58	1.54	1.17	1.65	1.34	1.40	1.33
4	1.84	1.82	1.33	1.93	1.52	1.61	1.51
5	1.90	1.89	1.35	1.99	1.54	1.62	1.52
6	1.86	1.85	1.26	1.94	1.45	1.56	1.44
7	2.00	2.01	1.43	2.15	1.66	1.74	1.63
8	1.97	1.99	1.42	2.09	1.62	1.72	1.61
9	1.62	1.66	1.07	1.74	1.25	1.35	1.24
10	1.37	1.39	0.80	1.43	0.95	1.03	0.92

Nonlinear Factor Loading Function - $G_\beta(X)$

- Fourth order Legendre polynomials to capture nonlinearities
  - $f(x) \approx 1 + x + \frac{1}{2}(3x^2 - 1) + \frac{1}{2}(5x^3 - 3x) + \frac{1}{8}(35x^4 - 30x^2 + 5)$

# Eigenvectors	CAPM	FF3	FF3+UMD	FF5	FF5+UMD	HXZ4	HXZ4+UMD
1	2.31	2.28	1.89	2.36	2.05	2.11	2.04
2	2.76	2.75	2.15	2.78	2.30	2.44	2.32
3	2.54	2.51	2.07	2.57	2.21	2.33	2.24
4	2.78	2.75	2.20	2.76	2.31	2.45	2.34
5	2.82	2.78	2.20	2.80	2.32	2.44	2.33
6	2.94	2.93	2.32	2.96	2.46	2.60	2.49
7	3.01	3.02	2.41	3.01	2.52	2.64	2.53
8	2.91	2.93	2.27	2.92	2.38	2.49	2.37
9	2.76	2.76	2.13	2.74	2.23	2.32	2.21
10	2.49	2.48	1.91	2.44	1.98	2.02	1.92

- Slightly higher alphas

Comparison with other "Characteristic Aggregators"

- Lewellen (2015) - rolling Fama-MacBeth slopes
- Stambaugh, Yuan (2016) - aggregate rank

	APC				RP-PC				IPCA			
	1 PC	5 PCs	10 PCs	20 PCs	1 RP-PC	5 RP-PCs	10 RP-PCs	20 RP-PCs	1 KPS-PC	5 KPS-PCs	10 KPS-PCs	20 KPS-PCs
Panel A: Arbitrage Portfolio												
alpha	2.45***	2.08***	1.28***	1.12***	2.45***	1.15***	1.06***	1.01***	2.39***	2.31***	2.15***	1.57***
	(0.29)	(0.25)	(0.26)	(0.25)	(0.29)	(0.30)	(0.27)	(0.26)	(0.25)	(0.26)	(0.42)	(0.33)
Adj. R^2	-0.00	0.29	0.34	0.37	-0.00	0.31	0.34	0.36	0.02	0.13	0.14	0.23
Num. obs.	612	612	612	612	612	612	612	612	612	612	612	612
Panel B: Lewellen												
alpha	3.01***	2.25***	0.50*	0.58***	3.01***	0.38	0.14	0.49*	3.00***	3.06***	2.36***	0.85**
	(0.31)	(0.20)	(0.27)	(0.22)	(0.31)	(0.28)	(0.28)	(0.25)	(0.29)	(0.28)	(0.60)	(0.41)
Adj. R^2	0.01	0.61	0.72	0.75	0.01	0.68	0.71	0.75	0.03	0.19	0.25	0.46
Num. obs.	612	612	612	612	612	612	612	612	612	612	612	612
Panel C: Stambaugh/Yuan												
alpha	1.44***	0.90***	0.09	−0.05	1.44***	−0.29***	−0.09	−0.12*	1.46***	0.95***	0.62**	−0.17
	(0.20)	(0.08)	(0.07)	(0.06)	(0.20)	(0.09)	(0.08)	(0.07)	(0.18)	(0.18)	(0.25)	(0.16)
Adj. R^2	0.12	0.92	0.96	0.97	0.12	0.94	0.96	0.97	0.31	0.47	0.61	0.85
Num. obs.	612	612	612	612	612	612	612	612	612	612	612	612

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

- Stambaugh, Yuan and Lewellen method forecast both well, but not α
- Arbitrage portfolio remains significant

Outline

Conclusion

- New methodology to estimate mispricing/risk in factor models
  - Role of characteristics may change
- Characteristics contain information about
  - factor loadings
  - mispricing
- Perfoms well in simulated economy
- Economically and statistically large alphas
- Extensions
  - ML technique for estimating G_α and G_β
  - International Market

Equivalent Estimation of Arbitrage Portfolio

Back

- Re-arrange

$$\overline{\mathsf{R}} = \mathsf{G}_{\alpha}(\mathsf{X}) + \Gamma_{\alpha} + (\mathsf{G}_{\beta}(\mathsf{X}) + \Gamma_{\beta})\overline{\mathsf{F}} + \overline{\mathsf{E}}$$

- to

$$\overline{\mathsf{R}} - \mathsf{G}_{\beta}(\mathsf{X})\overline{\mathsf{F}} = \mathsf{G}_{\alpha}(\mathsf{X}) + \left(\Gamma_{\alpha} + \Gamma_{\beta}\overline{\mathsf{F}}\right) + \overline{\mathsf{E}}$$

- Then regress

$$\overline{\mathsf{R}} - \mathsf{G}_{\beta}(\mathsf{X})\overline{\mathsf{F}} \text{ on } \mathsf{X}$$

Comparison with IPCA (KPS 2017, 2018)

IPCA

$$
\begin{aligned}
r_{i,t+1} &= \alpha_{i,t} + \beta_{i,t} f_{t+1} + \varepsilon_{i,t+1} \\
\alpha_{i,t} &= x_{i,t}'\theta + \nu_{\alpha,i,t} \\
\beta_{i,t} &= x_{i,t}'\gamma + \nu_{\beta,i,t}
\end{aligned}
$$

- Asymptotic setting $N, T \to \infty$
- Variation in α and β through variation in characteristics
- Relationship between characteristics and returns is fixed over time (θ,γ)
- Estimate α indirectly and iteratively

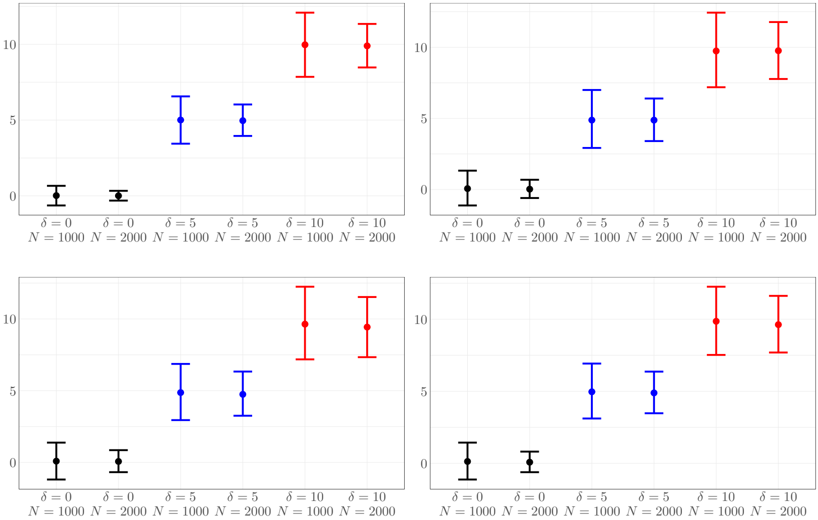
Comparison with IPCA (KPS 2017, 2018)

PPCA

$$
\begin{aligned}
r_{i,t+1} &= \alpha_i + \beta_i f_{t+1} + \varepsilon_{i,t+1} \\
\alpha_i &= x_i'\theta + \nu_{\alpha,i} \\
\beta_i &= x_i'\gamma + \nu_{\beta,i}
\end{aligned}
$$

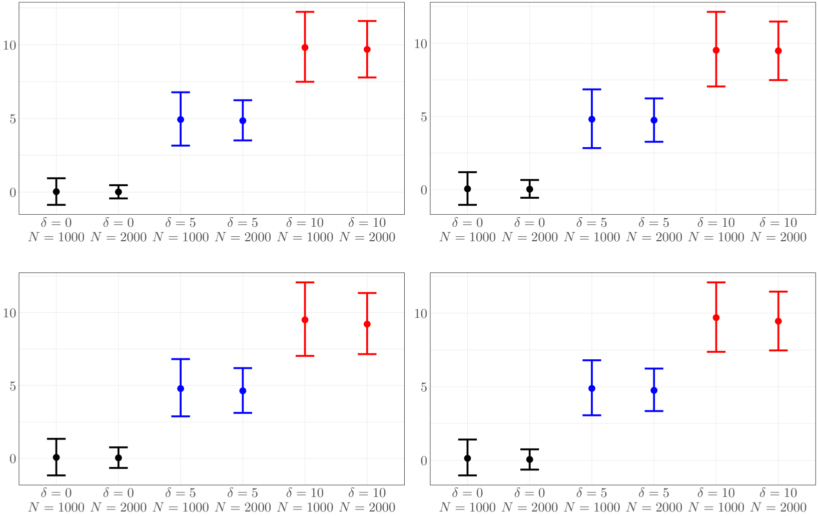
- Asymptotic setting $N \to \infty$, T fixed
- Variation in α and β through rolling estimation of θ and γ
- Relationship between characteristics and returns can change over time
    - learning
    - diminishing predictability
- Consistent estimator for α
    - Simulation evidence in favor of PPCA to estimate over short time windows

True Model Known



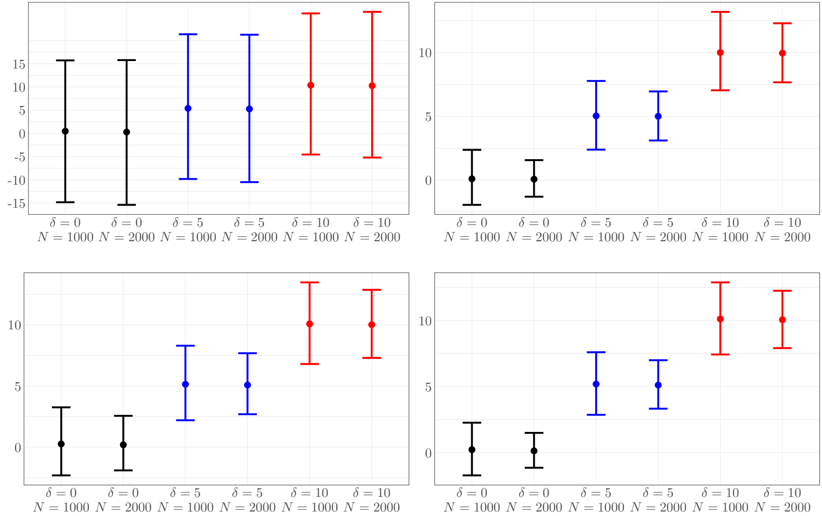
Back

Too Many Factors



Choose $K_{\text{wrong}} = K_{\text{true}} + 1$ Back

Too Few Factors



Choose $K_{\text{wrong}} = K_{\text{true}} - 1$ Back

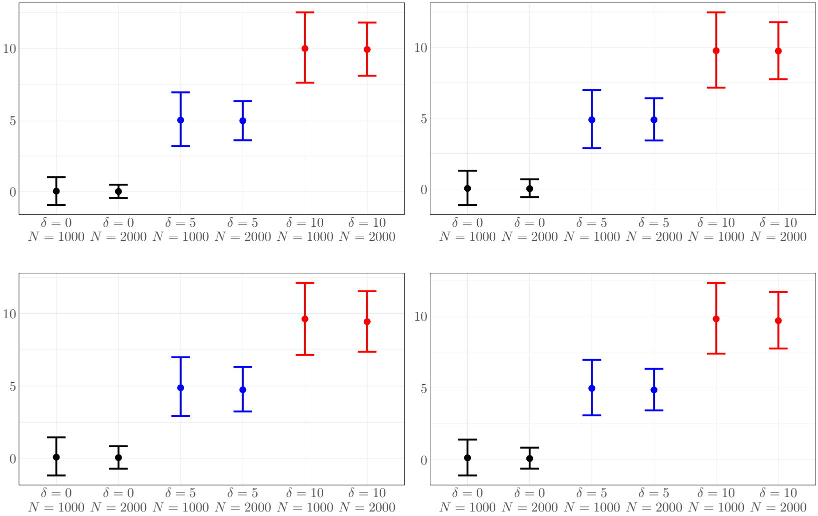
Simulation - Time-Varying Characteristics

- Reasonable robustness to non-constant characteristics
- Fix the initial characteristic over the calibration period as X
  - $x_{i,c}$ and $x_{i,c,t}$ denote the (i,c) element of X and X_t
- Generate X_t with $x_{i,c,t} = x_{i,c} + \rho_c \left(x_{i,c,t} - x_{i,c}\right) + \sigma_c \varepsilon_{i,t}$
  - ρ_c the estimated AR(1) coefficient of a certain characteristic c
  - σ_c^2 and variance of residuals of a certain characteristic c
  - $\varepsilon_{i,t}$ is drawn from $N(0,1)$ as i.i.d over i and t
- We then generate R_t, the t-th column of R, as follows:

$$\mathrm{R}_t = \boldsymbol{\alpha}_{t-1}\sqrt{\delta} + \mathrm{B}_{\mathrm{t}-1}\mathrm{f}_t + \mathrm{E}_t,$$

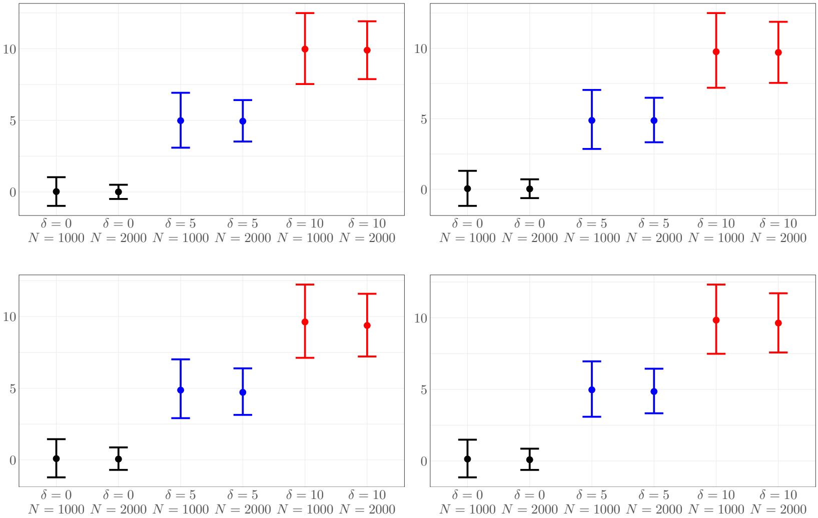
where $\boldsymbol{\alpha}_{t-1} = \mathrm{X}_{t-1}\boldsymbol{\theta}_{\boldsymbol{\alpha}}$, $\mathrm{B}_{t-1} = \mathrm{X}_{t-1}\boldsymbol{\Theta}_{\boldsymbol{\beta}} + \Gamma_\beta$, and E_t is the t-th column of E.

Simulation - Time-Varying Characteristics



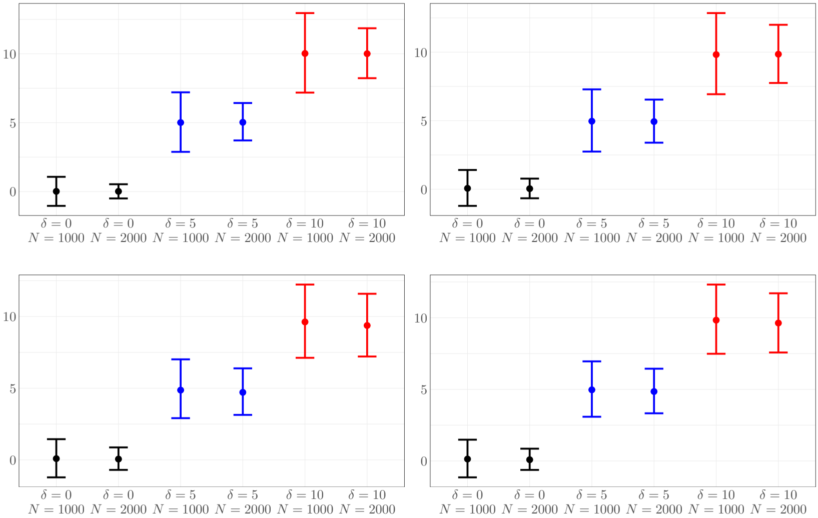
Results are not sensitive to AR(1) time-varying characteristics– Back

Simulation - Correlated Errors



Method is robust to industry clusters – Back

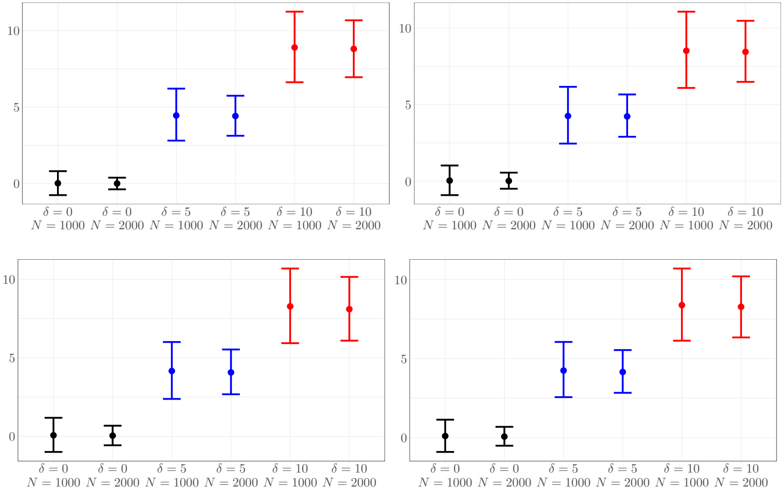
Simulation - Callibration (2006 - 2008)



More volatile callibration period – ▸ Back

Simulation - Omitted Characteristics

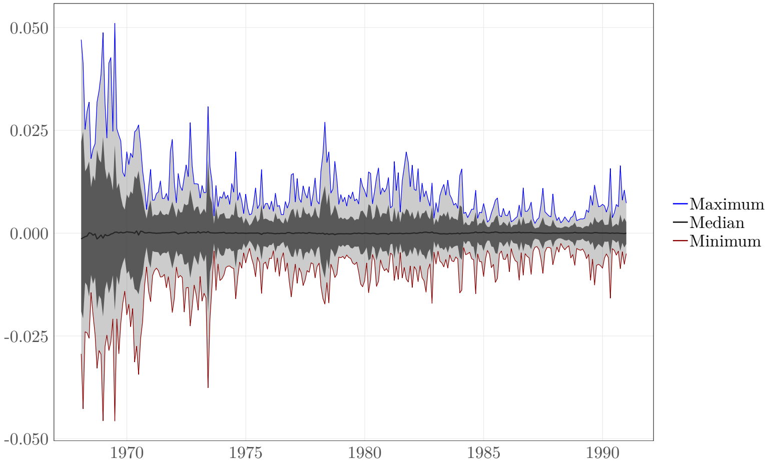
Back



- Drop randomly selected characteristics
- Slight underestimation of δ
- Omitting $x's$ does not mechanically introduce $\alpha > 0$, when the true $\alpha = 0$

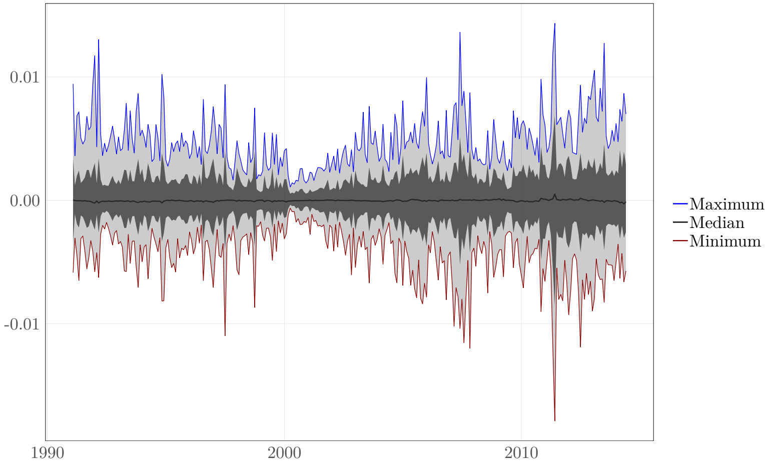
Arbitrage Portfolio - Portfolio Weights (I)

Portfolio weights from 1968 - 1990



Arbitrage Portfolio - Portfolio Weights (II)

Portfolio weights from 1991 - 2018



Arbitrage Portfolio - All Characteristics (▸ back)