Cross-sectional Predictability and Stock Returns

Svetlana Bryzgalova

Pasargad Advanced Research Initiative

Big Data Economics

Tehran, Iran

September 5, 2016

Fundamental theorem of asset pricing

- A stochastic discount factor is a stochastic process $\{M_{t,t+s}\}$ such that for any security with payoff x_{t+1} at time $t+1$ the price of that security at time t is

$$P_t = \mathbb{E}_t \left[M_{t,t+1} x_{t+1} \right]$$

- Equivalently,

$$1 = \mathbb{E} \left[M_{t,t+1}(1 + R_{t+1}) \right]$$

- The same pricing equation should hold for all the assets in the economy, including the risk-free rate:

$$\frac{1}{1 + R_{f,t+1}} = \mathbb{E} \left[M_{t,t+1} \right]$$

- Hence,

$$\mathbb{E}_t \left[M_{t,t+1}(R_{t+1} - R_{f,t+1}) \right] = \mathbb{E}_t \left[M_{t,t+1}(R^e_{t+1}) \right] = 0$$

Example: constructing an SDF

- Consider a one-period economy with $s = 1..S$ possible states of the world, each happening with a probability π_s.
- Arrow-Debreu securities: state-contingent claims that promise to pay 1 in a particular state of the world for the price of q_s today.
- Intuition for AD securities: basis in the space of payoff vectors.
- Under **law of one price (i.e. no arbitrage)**, the price of any security today that promises a stream of $\{x_s\}_{s=1}^{S}$ payoffs, depending on the state of the world tomorrow, is

$$P(x) = \sum_{s=1}^{S} q_s x_s = \sum_{s=1}^{S} \pi_s \frac{q_s}{\pi_s} x_s$$

- Define SDF as $m_s = \frac{q_s}{\pi_s}$. Then

$$P(x) = \sum_{s=1}^{S} \pi_s m_s x_s = \mathbb{E}\left[mx\right]$$

Fundamental theorem of asset pricing

Harrison and Kreps (1979), Hansen and Richard (1987):

- In complete markets under no arbitrage there exists a unique SDF that prices all the assets in the economy.
- Under imcomplete markets under no arbitrage, there exist multiple SDF that price all the assets in the economy.

Note, the general result is applied to multi-period economies, continuum of states, etc...

Asset returns are determined by their exposure to the pricing kernel and the price of risk:

$$\mathbb{E}\left[M_{t,t+1} r^e_{t+1}\right] = 0$$

$$\mathbb{E}\left[R^e_{t+1}\right] = -\frac{cov(M_{t,t+1}, R^e_{t+1})}{\mathbb{E}\left[M_{t,t+1}\right]} = \frac{cov(M_{t,t+1}, R^e_{t+1})}{var(M_{t,t+1})} \times \left(-\frac{var(M_{t,t+1})}{\mathbb{E}\left[M_{t,t+1}\right]}\right) = \beta \times \lambda_M$$

Any asset pricing model is tested on whether it can explain the **cross-section of asset returns**

Typical way of estimating: GMM or Fama-MacBeth regressions.

The Capital Asset Pricing Model (CAPM)

- According to the CAPM, there is only one source of risk: **Market risk**
- Investors are compensated for exposure to undiversifiable market risk
- Only market risk matters for expected returns
- CAPM equation:

$$\mathbb{E}[R_{i,t} - R_f] = \beta_i \mathbb{E}[R_{m,t} - R_f]$$

  where

$$\beta_i = \frac{Cov(R_{i,t}, R_{m,t})}{Var(R_{m,t})}$$

- The CAPM has two dimensions:
  - Time series given an asset i
  - Cross-section: Do assets with different β's have different excess returns?

CAPM: cross-section and time series

- The CAPM
$$\mathbb{E}[R_{i,t} - R_f] = \beta_i \mathbb{E}[R_{m,t} - R_f]$$
can be written as a linear regression:
$$R_{it} - R_f = \alpha_i + \beta_i (R_{mt} - R_f) + \epsilon_{i,t}$$
where
$$Cov(R_{m,t}, \epsilon_{i,t}) = 0$$
- α_i is called the **pricing error**
If the CAPM is true:
$$\alpha_i = 0$$
- Note: The CAPM should hold for **any** asset!
- *Only* market risk measured by β determines an asset's risk premium
- There are many asset characteristics that are associated with higher returns **for stocks with the same betas**.
- This started a quest for the right SDF, reflecting different dimensions of risk, as well as portfolios/types of securities that present a challenge.

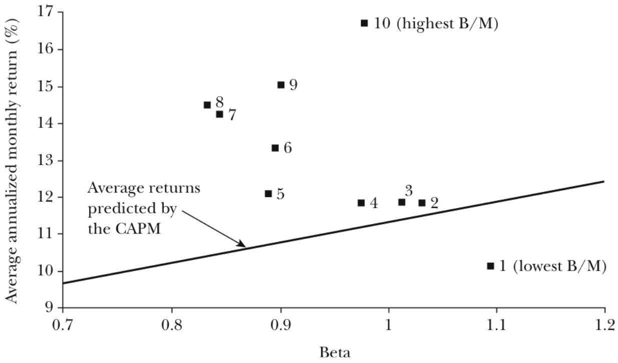
The demise of the CAPM: Value-growth portfolios

- Fama and French (1992)
- Standard measure of value/growth: A firm's book-to-market ration (B/M)
- Low B/M $\Rightarrow$ High market value relative to book value $\Rightarrow$ **Growth stock**
- High B/M $\Rightarrow$ low market value relative to book value $\Rightarrow$ **Value stock**
- Every June, sort firms according to their B/M and form portfolios; compute monthly portfolio returns From July to following June; resort according to current B/M and form new portfolios
- Data source: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`
- Let's start with 10 B/M portfolios

The demise of the CAPM: Value-growth portfolios

Figure 3

Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on B/M, 1963–2003



The demise of the CAPM: Size within value portfolios

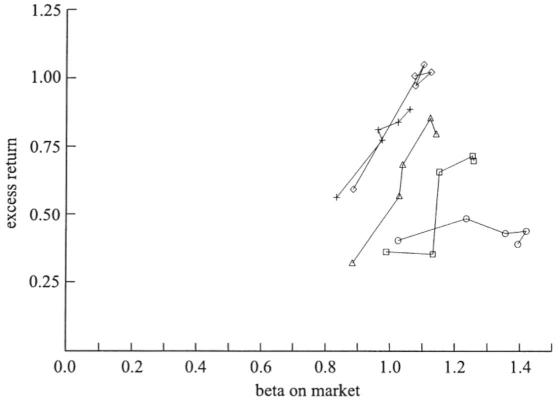


Figure 20.10. *Average excess returns vs. market beta. Lines connect portfolios with different size category within book market categories.*

The demise of the CAPM: Value within size portfolios

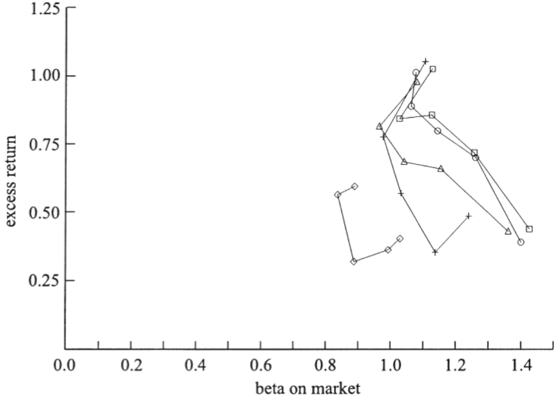


Figure 20.11. *Average excess returns vs. market beta. Lines connect portfolios with different book market categories within size categories.*

Formal tests

Recall that α_i should be zero:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i(R_{m,t} - R_{m,t}) + \epsilon_{i,t}$$

	Low	2	3	4	High	$H-L$	$t(H-L)$
Size-B/M Portfolios							
Small	0.73	1.32	1.36	1.57	1.67	0.59	4.13
2	0.89	1.15	1.40	1.45	1.55	0.48	3.62
3	0.90	1.22	1.20	1.35	1.51	0.37	2.64
4	1.01	0.99	1.22	1.34	1.37	0.36	2.75
Big	0.90	0.97	0.98	1.05	1.06	0.13	1.01
$S-B$	−0.14	0.26	0.28	0.31	0.39	0.38	3.32
$t(S-B)$	−0.77	1.46	1.85	2.18	2.63		
Size-E/P Portfolios							
Small	1.08	1.30	1.43	1.52	1.71	0.43	4.20
2	1.07	1.31	1.34	1.36	1.53	0.26	2.00
3	0.96	1.17	1.28	1.28	1.51	0.33	2.50
4	0.94	1.04	1.15	1.34	1.42	0.38	3.03
Big	0.85	0.95	0.92	1.19	1.13	0.26	2.07
$S-B$	0.18	0.31	0.34	0.17	0.35	0.33	3.19
$t(S-B)$	1.05	2.04	2.36	1.33	2.54		

Note: Fama and French (Journal of Finance, 1992), units are % per month.

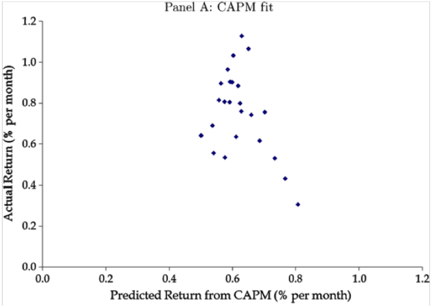
Cross-sectional regression with characteristics

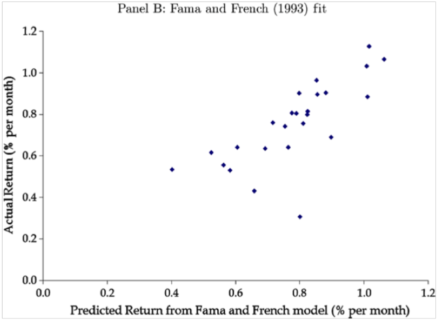
Include ME and B/M in cross-sectional regression. What does the CAPM predict?

$$R_{i,t} - R_{f,t} = b_0 + b_1 \hat{\beta}_{it} + b_2 \log\left(ME_{it}\right) + b_3 \log\left(BE/ME_{it}\right) + \alpha_{it}$$

β	ln(ME)	ln(BE/ME)	ln(A/ME)	ln(A/BE)	E/P Dummy	E(+)/P
0.15						
(0.46)						
	−0.15					
	(−2.58)					
−0.37	−0.17					
(−1.21)	(−3.41)					
		0.50				
		(5.71)				
			0.50	−0.57		
			(5.69)	(−5.34)		
					0.57	4.72
					(2.28)	(4.57)
	−0.11	0.35				
	(−1.99)	(4.44)				
	−0.11		0.35	−0.50		
	(−2.06)		(4.32)	(−4.56)		
	−0.16				0.06	2.99
	(−3.06)				(0.38)	(3.04)
	−0.13	0.33			−0.14	0.87
	(−2.47)	(4.46)			(−0.90)	(1.23)
	−0.13		0.32	−0.46	−0.08	1.15
	(−2.47)		(4.28)	(−4.45)	(−0.56)	(1.57)

Testing the FF 3-model





The next blow: Short-term momentum

Sort stocks according to returns over past 12 to 2 months.

Decile	Mean	Std. dev.	Alpha_1	Alpha_3
Panel B: Prior [−12 −2] return sorted portfolios				
Losers = 1	0.24	7.46	−0.92 (−5.74)	−1.03 (−6.60)
2	0.66	5.84	−0.37 (−3.34)	−0.48 (−4.34)
3	0.81	5.03	−0.13 (−1.36)	−0.23 (−2.48)
4	0.86	4.59	−0.05 (−0.69)	−0.14 (−2.01)
5	0.88	4.29	−0.01 (−0.15)	−0.10 (−1.68)
6	0.93	4.36	0.02 (0.39)	−0.04 (−0.79)
7	0.99	4.31	0.10 (1.57)	0.05 (0.83)
8	1.15	4.43	0.25 (4.00)	0.21 (3.47)
9	1.18	4.75	0.24 (3.48)	0.23 (3.30)
Winners = 10	1.56	5.94	0.52 (4.73)	0.62 (5.94)
Winners–losers	1.32	6.41	1.44 (6.29)	1.64 (7.16)

- W-L earns 1.32%*12=15.84% annually
- CAPM α is 1.44%*12=17.28%

Strategy summary

Data: monthly returns from 1932-2012

	MKT	SMB	HML	WML
Mean	7.48%	2.52%	4.83%	9.47%
Std. Dev.	19.15%	11.73%	12.57%	16.50%
Sharpe-ratio	0.39	0.21	0.38	0.57

- These returns are net of trading costs
- Momentum has very high turnover
- Short side of strategies hard to implement
- Many stocks involved are small and micro-caps

Fama-French multifactor models

Fama-French (1993):

$$\mathbb{E}\left[R_{i,t} - R_{f,t}\right] = \beta_{i,mkt}\mathbb{E}[R_{m,t} - R_{f,t}] + \beta_{i,smb}\mathbb{E}[\text{SMB}] + \beta_{i,hml}\mathbb{E}[\text{HML}]$$

- Fama and French argue that SMB and HML represent undiversifiable risk factors
- $\beta_{i,smb}$ and $\beta_{i,hml}$ measure the exposure of asset i to these risk factors
- The interpretation of these factors is (still) hotly debated
- Issues:
  - No theoretical foundation
  - FF do not explain why SMB and HML should be risk factors
  - What is the underlying economic reason that give rise to SMB and HML?

Testing the FF 3-model I

Data: 25 B/M-size sorted portfolio from Ken French's website, sample 1932-2012
Average returns; columns: growth to value; rows: small to large

mean returns

	low	2	3	4	high
small	0.6754	0.9987	1.2590	1.4377	1.5694
2	0.6995	1.0438	1.1844	1.2583	1.3534
3	0.8057	0.9712	1.0257	1.1405	1.2854
3	0.7366	0.8101	1.0062	1.0768	1.2294
large	0.6550	0.6441	0.8135	0.9058	1.0989

Comments on Fama-French model

- The precise meaning of the FF model is (still) hotly debated
- It is agreed that the CAPM is dead and the FF model produces much smaller pricing errors (even though the FF model is statistically rejected)
- But are HML and SMB true risk factors?
- Fama-French: Yes, they are
- Others are more skeptical

- My view: HML and SMB are summaries of the value and size puzzles but they are not explanations of the puzzles.

  Indeed, they should be left-hand-side variables, i.e. portfolios to be explained.

  However, the FF model is useful in practice as a reduced-form model.

Risk models in practice

Richardson, Tuna and Wysocki (2010): Survey of 201 investment managers and 63 academics

Q1: Which risk model is most appropriate for risk calibration of an equity trading strategy?

	Practitioner Opinions	*Academic Opinions*
CAPM with size & industry adjustments	35%	*7%* **
Fama-French 3-factor model (Market, Size, Book Value/Market Value)	24%	*22%*
Multifactor model	11%	*4%* **
Other model	11%	*15%*
CAPM	10%	*4%* *
Fama-French 3-factor model plus other factors	5%	*33%* **
CAPM with size adjustments	4%	*15%* **

* and ** indicate difference in means across practitioner and academic sample answers are significant at 5 and 1% levels, respectively.

50 years of empirical asset pricing in a nutshell

- The real implication of any asset pricing model is not how much of the returns time series in can explain, but how well it handles the cross-section of asset returns.
- Differences in exposure to systematic risk should justify differences in risk premia across various assets
- Throughout the years there has been accumulated evidence for a variety of factors being "priced".

50 years of empirical asset pricing in a nutshell

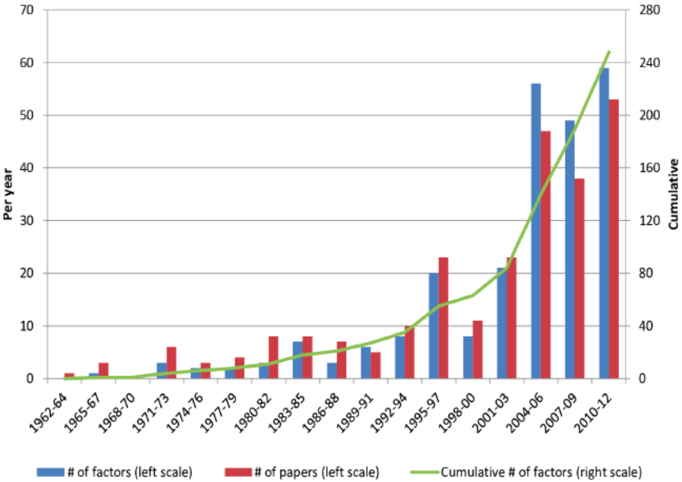
- The real implication of any asset pricing model is not how much of the returns time series in can explain, but how well it handles the cross-section of asset returns.
- Differences in exposure to systematic risk should justify differences in risk premia across various assets
- Throughout the years there has been accumulated evidence for a variety of factors being "priced".
- Why so many?

What prices the cross-section of stock returns?

Factor classification for the cross-section of stock returns, Harvey, Liu and Zhu (2016)

Risk classification		Description	Examples
Common (113)	**Financial** (46)	Proxy for aggregate financial market movement, including market portfolio returns, volatility, squared market returns, etc.	Sharpe (1964): market returns; Kraus and Litzenberger (1976): squared market returns
	Macro (40)	Proxy for movement in macroeconomic fundamentals, including consumption, investment, inflation, etc.	Breeden (1979): consumption growth; Cochrane (1991): investment returns
	Microstructure (11)	Proxy for aggregate movements in market microstructure or financial market frictions, including liquidity, transaction costs, etc.	Pastor and Stambaugh (2003): market liquidity; Lo and Wang (2006): market trading volume
	Behavioral (3)	Proxy for aggregate movements in investor behavior, sentiment or behavior-driven systematic mispricing	Baker and Wurgler (2006): investor sentiment; Hirshleifer and Jiang (2010): market mispricing
	Accounting (8)	Proxy for aggregate movement in firm-level accounting variables, including payout yield, cash flow, etc.	Fama and French (1992): size and book-to-market; Da and Warachka (2009): cash flow
	Other (5)	Proxy for aggregate movements that do not fall into the above categories, including momentum, investors beliefs, etc.	Carhart (1997): return momentum; Ozoguz (2008): investors beliefs
Individual (202)	**Financial (61)**	Proxy for firm-level idiosyncratic financial risks, including volatility, extreme returns, etc.	Ang, Hodrick, Xing and Zhang (2006): idiosyncratic volatility; Bali, Cakici and Whitelaw (2011): extreme stock returns
	Microstructure (28)	Proxy for firm-level financial market frictions, including short sale restrictions, transaction costs, etc.	Jarrow (1980): short sale restrictions; Mayshar (1981): transaction costs
	Behavioral (3)	Proxy for firm-level behavioral biases, including analyst dispersion, media coverage, etc.	Diether, Malloy and Scherbina (2002): analyst dispersion; Fang and Peress (2009): media coverage
	Accounting (86)	Proxy for firm-level accounting variables, including PE ratio, debt to equity ratio, etc.	Basu (1977): PE ratio; Bhandari (1988): debt to equity ratio
	Other (24)	Proxy for firm-level variables that do not fall into the above categories, including political campaign contributions, ranking-related firm intangibles, etc.	Cooper, Gulen and Ovtchinnikov (2010): political campaign contributions; Edmans (2011): intangibles

Factor production mill



The universe of factors

- Harvey, Liu and Zhu (2016): data mining, publication bias, multiple testing

The universe of factors

- Harvey, Liu and Zhu (2016): data mining, publication bias, multiple testing
- Problem with the inference?

Asset pricing with spurious factors

Linear factor model:

$$\textit{Expected Return} = \textit{risk} \ \times \textit{risk premium}$$

Asset pricing with spurious factors

Linear factor model:

$$\textit{Expected Return} = \textit{risk} \times \textit{risk premium}$$

Example: Capital Asset Pricing Model (Sharpe (1961), Lintner (1965))

$$E(R - r_f) = \underbrace{\mathbf{0_n}}_{\lambda_0} + \beta_{mkt} \underbrace{E(R_{mkt} - r_f)}_{\lambda_{0,F}}$$

Asset pricing with spurious factors

Linear factor model:

$$\textit{Expected Return} = \textit{risk} \times \textit{risk premium}$$

Example: Capital Asset Pricing Model (Sharpe (1961), Lintner (1965))

$$E(R - r_f) = \underbrace{\mathbf{0_n}}_{\lambda_0} + \beta_{mkt} \underbrace{E(R_{mkt} - r_f)}_{\lambda_{0,F}}$$

General setting:

$$E(R_t^e) = i_n \lambda_{0,c} + \beta_F \lambda_{0,F}$$
$$cov(R_t^e, F_t) = \beta_F var(F_t)$$

Asset pricing with spurious factors

Linear factor model:

$$Expected\ Return = risk \times risk\ premium$$

Example: Capital Asset Pricing Model (Sharpe (1961), Lintner (1965))

$$E(R - r_f) = \underbrace{\mathbf{0_n}}_{\lambda_0} + \beta_{mkt} \underbrace{E(R_{mkt} - r_f)}_{\lambda_{0,F}}$$

General setting:

$$E(R_t^e) = i_n \lambda_{0,c} + \beta_F \lambda_{0,F}$$
$$cov(R_t^e, F_t) = \beta_F var(F_t)$$

Typical approach: Fama-MacBeth two-pass procedure (or GMM)

- time series regression of excess returns on factors: $\widehat{\beta}_F$
- cross-sectional regression of average excess returns $\bar{R}^e$ on betas $\hat{\beta} = [i \quad \hat{\beta}_F]$:
$$\hat{\lambda}_{OLS} = \left[\hat{\beta}' \hat{\beta}\right]^{-1} \hat{\beta}' \bar{R}^e$$

Asset pricing with spurious factors

Linear factor model:

$$\textit{Expected Return} = \textit{risk} \times \textit{risk premium}$$

Example: Capital Asset Pricing Model (Sharpe (1961), Lintner (1965))

$$E(R - r_f) = \underbrace{\mathbf{0_n}}_{\lambda_0} + \beta_{mkt} \underbrace{E(R_{mkt} - r_f)}_{\lambda_{0,F}}$$

General setting:

$$E(R_t^e) = i_n \lambda_{0,c} + \beta_F \lambda_{0,F}$$
$$cov(R_t^e, F_t) = \beta_F var(F_t)$$

Typical approach: Fama-MacBeth two-pass procedure (or GMM)

- time series regression of excess returns on factors: $\widehat{\beta}_F$
- cross-sectional regression of average excess returns $\bar{R}^e$ on betas $\hat{\beta} = [i \;\; \hat{\beta}_F]$:
$$\hat{\lambda}_{OLS} = \left[\hat{\beta}' \hat{\beta}\right]^{-1} \hat{\beta}' \bar{R}^e$$

Kan and Zhang (1999ab), Kleibergen (2009), Burnside (2014): If a factor only weakly correlates with asset returns, $\beta_j = \frac{B}{\sqrt{T}}$ (or even $\beta_j = \mathbf{0}_{n \times 1}$), standard estimation techniques fail.

Illustrative experiment

- 25 Fama-French portfolios (1947Q2 - 2014Q2)
- Simulate a normal random variable with the same mean and variance as nondurable consumption growth
- Estimate a 4-factor model (market, size, book-to-market + the spurious factor)
- Repeat 1000 times

Spurious factor **"priced"** at 10% significance level in the cross-section of stocks:

1. Nontradable factor:
    - Fama-MacBeth with OLS/HC standard errors:
    - Fama-MacBeth with Shanken standard errors:

Illustrative experiment

- 25 Fama-French portfolios (1947Q2 - 2014Q2)
- Simulate a normal random variable with the same mean and variance as nondurable consumption growth
- Estimate a 4-factor model (market, size, book-to-market + the spurious factor)
- Repeat 1000 times

Spurious factor **"priced"** at 10% significance level in the cross-section of stocks:

1. Nontradable factor:
   - Fama-MacBeth with OLS/HC standard errors: 57.5%
   - Fama-MacBeth with Shanken standard errors: 48.1%
2. Tradable mimicking portfolio:
   - Fama-MacBeth with OLS/HC standard errors:
   - Fama-MacBeth with Shanken standard errors:

Illustrative experiment

- 25 Fama-French portfolios (1947Q2 - 2014Q2)
- Simulate a normal random variable with the same mean and variance as nondurable consumption growth
- Estimate a 4-factor model (market, size, book-to-market + the spurious factor)
- Repeat 1000 times

Spurious factor **"priced"** at 10% significance level in the cross-section of stocks:

1. Nontradable factor:
   - Fama-MacBeth with OLS/HC standard errors: 57.5%
   - Fama-MacBeth with Shanken standard errors: 48.1%
2. Tradable mimicking portfolio:
   - Fama-MacBeth with OLS/HC standard errors: 69.2%
   - Fama-MacBeth with Shanken standard errors: 49.2%

Illustrative experiment

- 25 Fama-French portfolios (1947Q2 - 2014Q2)
- Simulate a normal random variable with the same mean and variance as nondurable consumption growth
- Estimate a 4-factor model (market, size, book-to-market + the spurious factor)
- Repeat 1000 times

Spurious factor **"priced"** at 10% significance level in the cross-section of stocks:

1. Nontradable factor:
   - Fama-MacBeth with OLS/HC standard errors: 57.5%
   - Fama-MacBeth with Shanken standard errors: 48.1%
2. Tradable mimicking portfolio:
   - Fama-MacBeth with OLS/HC standard errors: 69.2%
   - Fama-MacBeth with Shanken standard errors: 49.2%

GLS, GMM, CU-GMM, etc: same problem under the suitable normalisation

Generalised Method of Moments

GMM

$$\hat{\theta} = \arg\min_{\theta \in S} \left[\frac{1}{T} \sum_{t=1}^{T} g_t(\theta) \right]' W_T \left[\frac{1}{T} \sum_{t=1}^{T} g_t(\theta) \right]$$

where $W_T(\theta)$ is a p.d. weight $(n + nk + k) \times (n + nk + k)$ matrix, and

$$g_t(\theta) = \begin{bmatrix} R_t - i_n\lambda_c - \beta(\lambda_F - \mu + F_t) \\ vec\left([R_t - i_n\lambda_c - \beta(\lambda_F - \mu + F_t)]\, F_t'\right) \\ F_t - \mu \end{bmatrix}$$

is a sample moment of the dimension $(n + nk + k) \times 1$.

In the presence of a useless factor, the model is not identified, since $G(\theta_0) = E[G_t(\theta_0)] = E\left[\frac{dg_t(\theta_0)}{d\theta}\right]$ will have a reduced column rank.

Any nonlinear model admitting a beta-representation could have the same problem.

Spurious factors → lack of identification

E.g. Kleibergen (2009), Kleibergen and Zhou (2013), Gospodinov, Kan and Robotti (2014a,b)

Correctly specified model: $E(R^e_{i,t}) = \lambda_0 + \beta_{i,f}\lambda_{0,f}$

- $\hat{\lambda}$ for strong factors are consistent, but highly non-normal and varies a lot.
- $\hat{\lambda}$ for spurious factors converge to a random variable
- R^2, $GLS\ R^2$ tend to be inflated, and follow non-standard distributions
- Hansen-Jagannathan test for correct model specification is invalid

Misspecified model: $E(R^e_t) = \lambda_{i,0} + \beta_{i,f}\lambda_{0,f}$

- $\hat{\lambda}$ for strong factors are inconsistent
- $\hat{\lambda}$ for spurious factors diverge with the sample size
- $t - stat$ for spurious factors tend to infinity
- R^2, $GLS\ R^2$ are substantially inflated, and follow non-standard distributions
- HJ test for correct model specification is invalid
- The problems are exacerbated when the set of testing portfolios is large

A solution: Pen-FM estimator

Modified Fama-MacBeth procedure

- time series regression of excess returns on factors: $\widehat{\beta}_F$
- Risk premia estimates minimise a penalised version of the 2nd stage:
  $\hat{\lambda} = \arg\min_{\lambda \in M} \frac{1}{2N} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right]$

A solution: Pen-FM estimator

Modified Fama-MacBeth procedure

- time series regression of excess returns on factors: $\widehat{\beta}_F$
- Risk premia estimates minimise a penalised version of the 2nd stage:
  $\hat{\lambda} = \underset{\lambda \in M}{\arg\min} \frac{1}{2N} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right] + \hat{\sigma} T^{-d/2} \sum_{j=1}^{k} w_j \left|\lambda_j\right|$

where $w_j = \left(\sum_{i=1}^{N} |z(\hat{\rho}_{ij})|\right)^{-d}$, $\hat{\sigma}^2 = \hat{var}(\epsilon_{i,t})$, $d > 2$, $\hat{\rho_{i,j}}$ is the partial correlation between portfolio i and factor j, and $z(\cdot)$ is Fisher's z-transformation.

A solution: Pen-FM estimator

Modified Fama-MacBeth procedure

- time series regression of excess returns on factors: $\widehat{\beta}_F$
- Risk premia estimates minimise a penalised version of the 2nd stage:
  $\hat{\lambda} = \arg\min_{\lambda \in M} \frac{1}{2N} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right] + \hat{\sigma} T^{-d/2} \sum_{j=1}^{k} w_j \left|\lambda_j\right|$

where $w_j = \left(\sum_{i=1}^{N} |z(\hat{\rho}_{ij})|\right)^{-d}$, $\hat{\sigma}^2 = \hat{var}(\epsilon_{i,t})$, $d > 2$, $\hat{\rho_{i,j}}$ is the partial correlation between portfolio i and factor j, and $z(\cdot)$ is Fisher's z-transformation.

- The penalty is inversely proportional to the total strength of the factor for a given set of portfolios:
  - for a spurious factor, $\hat{\rho}_{ij} \xrightarrow{p} 0$, hence $L1-$ norm of $z(\hat{\rho}_{ij})$ is small, $O(\frac{1}{\sqrt{T}})$
  - for a strong factor, $\hat{\rho}_{ij} \xrightarrow{p} const$, hence $L1-$ norm of $z(\hat{\rho}_{ij})$ is $O(1)$
  - spurious factor risk premia is picked by the penalty term and set to 0

Robust to simple data scaling. Betas, partial correlations, t-stats can also be used

Shrinkage estimators: A tool for model selection

- **Pen-FM**: penalize factors, depending on the nature of β_j (whether their impact is identified)

$$\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}^k} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right] + \eta_T \sum_{j=1}^{n} \mathbf{w_j} |\lambda_\mathbf{j}|$$

where $\mathbf{w_j} = \left(\sum_{\mathbf{i=1}}^{\mathbf{N}} |\mathbf{z}(\hat{\rho}_\mathbf{ij})|\right)^{-\mathbf{d}}$, $d > 2$

- **LASSO** (Least Absolute Shrinkage and Selection Operator), Tibshirani (1996)

$$\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}^k} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right] + \eta_T \sum_{j=1}^{n} \mathbf{1} |\lambda_\mathbf{j}|$$

- **Adaptive LASSO**, Zou (2006): penalize factors inversely proportionally to their effect on Y

$$\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}^k} \left[\bar{R}^e - \hat{\beta}\lambda\right]' \left[\bar{R}^e - \hat{\beta}\lambda\right] + \eta_T \sum_{j=1}^{n} \mathbf{w_j} |\lambda_\mathbf{j}|$$

where $\mathbf{w_j} = \frac{\mathbf{1}}{|\hat{\lambda}_\mathbf{j}^\mathbf{ols}|^\mathbf{d}}$, $d > 0$.

Pen-FM: Asymptotic distribution

Lemma

Under Assumption A1, average cross-sectional returns and OLS estimator $\hat{\beta}$ have a joint large sample distribution:

$$\sqrt{T}\begin{pmatrix} \bar{R} - \beta\lambda_f \\ vec(\hat{\beta} - \beta) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_R \\ \psi_\beta \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega & 0 \\ 0 & V_{ff}^{-1} \otimes \Omega \end{pmatrix}\right]$$

where ψ_R is independent of $\psi_\beta = (V_{ff}^{-1} \otimes I_n)(\varphi_\beta - (\mu_f \otimes I_n)\psi_R)$

Theorem

Under the conditions of Lemma 1, if $W_T \xrightarrow{p} W$, W is a positive definite $n \times n$ matrix, $\eta_T = \eta T^{-d/2}$ with a finite constant $\eta > 0$, $d > 0$ and $\beta'_{ns}\beta_{ns}$ having full rank, $\hat{\lambda}_{ns} \xrightarrow{p} \lambda_{0,ns}$ and $\hat{\lambda}_{sp} \xrightarrow{p} 0$

Further, if $d > 2$

$$\sqrt{T}\begin{pmatrix} \hat{\lambda}_{ns} - \lambda_{0,ns} \\ \hat{\lambda}_{sp} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} [\beta'_{ns} W \beta_{ns}]^{-1} \beta'_{ns} W \Psi_{\beta,ns}\lambda_{0,ns} + (\beta'_{ns} W \beta_{ns})^{-1} \beta'_{ns} W \psi_R \\ 0 \end{pmatrix}$$

Horse Race

Simulation design of GKR (2014a): sequential procedure based on modified $t - statistic$ that eliminate both spurious and unpriced factors

- 25 size and book-to-market portfolios + 17 industry portfolios
- Monthly data
- Parameters from the estimated linear SDF model with 3 Fama-French factors
- Consider two settings for the estimation:
  - 2 useful factors ($\lambda \neq 0$), 1 unpriced factor ($\lambda = 0$), 1 useless factor
  - 2 useful factors ($\lambda \neq 0$), 2 useless factors.
- Model correctly or incorrectly specified

- Focus on the survival rates of various factors

Factor survival rates, Pen-FM vs t_m test of GKR (2014)

Survival rates of useful and irrelevant factors (based on 10 000 simulations)

Panel A: Correctly specified model

	Useful factor			Useful factor			Unpriced factor			Useless factor		
	$(\lambda_1 \neq 0)$			$(\lambda_2 \neq 0)$			$(\lambda_3 = 0)$					
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
200	0.5142			0.6599			0.0231			0.0023		
600	0.9864			0.9987			0.0141			0.0006		
1000	0.9999			1.0000			0.0117			0.0003		

	Useful factor			Useful factor			Useless factor			Useless factor		
	$(\lambda_1 \neq 0)$			$(\lambda_2 \neq 0)$								
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
200	0.5265			0.6582			0.0017			0.0025		
600	0.9880			0.9985			0.0007			0.0003		
1000	1			0.9900			0.0000			0.0002		

Factor survival rates, Pen-FM vs t_m test of GKR (2014)

Survival rates of useful and irrelevant factors (based on 10 000 simulations)

Panel A: Correctly specified model

T	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Unpriced factor ($\lambda_3 = 0$)			Useless factor		
	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
200	0.5142	1	0.9998	0.6599	0.9787	0.9686	0.0231	0.9788	0.9749	0.0023	0	0.0040
600	0.9864	1	0.9998	0.9987	0.9802	0.9784	0.0141	0.9777	0.9813	0.0006	0	0.0027
1000	0.9999	1	0.9999	1.0000	0.9828	0.9815	0.0117	0.9833	0.9829	0.0003	0	0.0023

T	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Useless factor			Useless factor		
	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
200	0.5265	1	1	0.6582	1	1	0.0017	0	0.0044	0.0025	0	0.0059
600	0.9880	1	1	0.9985	1	1	0.0007	0	0.0029	0.0003	0	0.0034
1000	1	1	1	0.9900	1	1	0.0000	0	0.0025	0.0002	0	0.0029

Factor survival rates, Pen-FM vs t_m test of GKR (2014)

Survival rates of useful and irrelevant factors (based on 10 000 simulations)

Panel A: Correctly specified model

	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Unpriced factor ($\lambda_3 = 0$)			Useless factor		
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
50	0.0628	1	0.9995	0.1120	0.9166	0.9245	0.0588	0.9233	0.9426	0.1541	0	0.0126
100	0.1760	1	0.9997	0.2431	0.9403	0.9439	0.0207	0.9539	0.9606	0.0072	0	0.0079
150	0.3444	1	0.9998	0.4623	0.9652	0.9598	0.0232	0.9622	0.9695	0.0031	0	0.0053
200	0.5142	1	0.9998	0.6599	0.9787	0.9686	0.0231	0.9788	0.9749	0.0023	0	0.0040
250	0.6614	1	0.9998	0.8035	0.9761	0.9742	0.0231	0.9746	0.9786	0.0017	0	0.0032
600	0.9864	1	0.9998	0.9987	0.9802	0.9784	0.0141	0.9777	0.9813	0.0006	0	0.0027
1000	0.9999	1	0.9999	1.0000	0.9828	0.9815	0.0117	0.9833	0.9829	0.0003	0	0.0023

	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Useless factor			Useless factor		
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
50	0.0573	1	0.9999	0.0984	1	1	0.1537	0	0.0164	0.1485	0	0.0232
100	0.1739	1	0.9999	0.2351	1	1	0.0068	0	0.0065	0.0085	0	0.0119
150	0.2020	1	1	0.2290	1	1	0.0080	0	0.0059	0.0032	0	0.0079
200	0.5265	1	1	0.6582	1	1	0.0017	0	0.0044	0.0025	0	0.0059
250	0.6742	1	1	0.8080	1	1	0.0015	0	0.0035	0.0015	0	0.0040
600	0.9880	1	1	0.9985	1	1	0.0007	0	0.0029	0.0003	0	0.0034
1000	1	1	1	0.9900	1	1	0.0000	0	0.0025	0.0002	0	0.0029

Factor survival rates, Pen-FM vs t_m test of GKR (2014)

Survival rates of useful and irrelevant factors (based on 10 000 simulations)

Panel B: Misspecified model

	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Unpriced factor ($\lambda_3 = 0$)			Useless factor		
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
50	0.0640	1	0.9995	0.1167	0.9453	0.9122	0.0676	0.9132	0.9437	0.1790	0	0.0133
100	0.1696	1	0.9996	0.2353	0.9617	0.9415	0.0224	0.9425	0.9614	0.0142	0	0.0075
150	0.3343	1	0.9997	0.4389	0.9733	0.9566	0.0221	0.9566	0.9699	0.0088	0	0.0052
200	0.5016	1	0.9998	0.6298	0.9787	0.9653	0.0240	0.9652	0.9751	0.0080	0	0.0039
250	0.6526	1	0.9998	0.7750	0.9826	0.9713	0.0238	0.9775	0.9786	0.0079	0	0.0031
600	0.9806	1	0.9998	0.9963	0.9850	0.9758	0.0138	0.9751	0.9812	0.0073	0	0.0026
1000	0.9972	1	0.9998	0.9989	0.9871	0.9792	0.0121	0.9764	0.9830	0.0088	0	0.0022

	Useful factor ($\lambda_1 \neq 0$)			Useful factor ($\lambda_2 \neq 0$)			Unpriced factor ($\lambda_3 = 0$)			Useless factor		
T	$t_m(\lambda_1)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_2)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_3)$	Pen-FM (pointwise)	Pen-FM (bootstrap)	$t_m(\lambda_4)$	Pen-FM (pointwise)	Pen-FM (bootstrap)
50	0.0406	1	0.9986	0.0815	0.7971	0.8105	0.1669	0	0.0228	0.1660	0	0.0279
100	0.0985	1	0.9992	0.1310	0.8352	0.8581	0.0138	0	0.0150	0.0141	0	0.0184
150	0.1928	1	0.9994	0.2493	0.8533	0.8634	0.0083	0	0.0123	0.0093	0	0.0134
200	0.3058	1	0.9996	0.3840	0.9071	0.8937	0.0074	0	0.0101	0.0081	0	0.0103
250	0.4221	1	0.9997	0.5180	0.8928	0.9027	0.0073	0	0.0082	0.0078	0	0.0087
600	0.9026	1	0.9997	0.9516	0.9204	0.9380	0.0097	0	0.0069	0.0086	0	0.0073
1000	0.9822	1	0.9997	0.9922	0.9628	0.9496	0.0102	0	0.0059	0.0096	0	0.0063

Pen vs Adaptive Lasso: factor survival rate

Comparison of the Pen-FM estimator with the adaptive lasso, based on the survival rates of useful and useless factors.

Panel A: Correctly specified model

	Useful ($\lambda_1 \neq 0$)		Useful ($\lambda_2 \neq 0$)		Useful ($\lambda_3 = 0$)		Useless	
T	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)
50	1	0.4172	0.9166	1	0.9233	0.7340	0	1
100	1	0.4745	0.9403	1	0.9539	0.8392	0	1
150	1	0.5173	0.9652	1	0.9622	0.9262	0	1
200	1	0.5743	0.9787	1	0.9748	0.9431	0	1
250	1	0.6260	0.9761	1	0.9746	0.9694	0	1
600	1	0.8132	0.9802	1	0.9777	1	0	1
1000	1	0.9099	0.9828	1	0.9833	1	0	1

	Useful ($\lambda_1 \neq 0$)		Useful ($\lambda_2 \neq 0$)		Useless		Useless	
T	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)
50	1	1	1	0.9322	0	1	0	1
100	1	1	1	0.9851	0	1	0	1
150	1	1	1	0.9955	0	1	0	1
200	1	1	1	1	0	1	0	1
250	1	1	1	1	0	1	0	1
600	1	1	1	1	0	1	0	1
1000	1	1	1	1	0	1	0	1

Pen vs Adaptive lasso: factor survival rate

Comparison of the Pen-FM estimator with the adaptive lasso, based on the survival rates of useful and useless factors.

Panel B: Misspecified model

	Useful ($\lambda_1 \neq 0$)		Useful ($\lambda_2 \neq 0$)		Useful ($\lambda_3 = 0$)		Useless	
T	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)
50	1	0.4691	0.9453	1	0.9132	0.6133	0	1
100	1	0.4782	0.9617	1	0.9424	0.7134	0	1
150	1	0.4784	0.9733	1	0.9566	0.7650	0	1
200	1	0.4870	0.9787	1	0.9652	0.7612	0	1
250	1	0.4566	0.9826	1	0.9775	0.8377	0	1
600	1	0.5179	0.9850	1	0.9751	0.9810	0	1
1000	1	0.6433	0.9989	1	0.9764	0.9959	0	1

	Useful ($\lambda_1 \neq 0$)		Useful ($\lambda_2 \neq 0$)		Useless		Useless	
T	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)	Pen-FM	AdaLasso (BIC)
50	1	0.5352	0.7971	1	0	0.5654	0	1
100	1	0.4833	0.8352	1	0	0.6132	0	1
150	1	0.5090	0.8553	1	0	0.6911	0	1
200	1	0.4566	0.9071	1	0	0.7177	0	1
250	1	0.3431	0.9121	1	0	0.7432	0	1
600	1	0.3217	0.9204	1	0	0.9210	0	1
1000	1	0.2918	0.9628	1	0	0.9618	0	1

Tuning parameters

Empirical applications

- Stocks: monthly / quarterly / yearly portfolios of stocks sorted by
  - size and book-to-market
  - industry / beta / volatility / past 12 month return,
  - asset growth / total accruals / stock issuance

Over 40 various specifications, including Fama-French factors, consumption growth rates, investment, etc.

Cross-section of stocks and tradable factors

Cross-section of stocks and tradable factors.

Factors	p-value (Wald)	t_m surv	Fama-MacBeth estimator λ_j	p-value (OLS)	p-value (Shanken)	p-value (Bootstrap)	R^2 (%)	Pen-FM estimator λ_j	Shrinkage rate (Bootstrap)	p-value (Bootstrap)	R^2 (%)
CAPM											
Intercept	-	-	1.431***	0.0008	0.0009	0.002	19	1.431***	0	0.002	19
MKT	0	yes	-0.658	0.1222	0.1674	0.184		-0.658	0	0.184	
Fama and French (1992)											
Intercept	-	-	1.252	0	0	0	70	1.2533	0	0	70
MKT	0	yes	-0.703*	0.0205	0.0587	0.06		-0.704*	0	0.06	
SMB	0	yes	0.145	0	0.3083	0.376		0.145	0	0.376	
HML	0	no	0.43***	0	0.0018	0.008		0.429***	0	0.008	
Asness, Frazzini and Pedersen (2014)											
Intercept	-	-	0.7*	0.0317	0.0409	0.092	84	0.576	0	0.212	83
MKT	0	yes	-0.327	0.3177	0.4155	0.412		-0.206	0	0.684	
SMB	0	yes	0.174	0	0.2325	0.288		0.172	0	0.292	
HML	0	no	0.398**	0	0.0041	0.016		0.416***	0	0.008	
QMJ	0	no	0.44**	0.0001	0.006	0.016		0.324*	0.084	0.084	

Cross-section of stocks and nontradable factors

Cross-section of stocks and non-tradable factors: Yogo (2006).

Factors	p-value (Wald)	t_m surv	Fama-MacBeth estimator λ_j	p-value (OLS)	p-value (Shanken)	p-value (Bootstrap)	R^2 (%)	Pen-FM estimator λ_j	Shrinkage rate (Bootstrap)	p-value (Bootstrap)	R^2 (%)
25 Fama-French portfolios											
Intercept	-	-	2.335**	0.0123	0.1209	0.03	55	3.445***	0	0.002	11
Nondur	0.1116	no	0.641	0.0035	0.0721	0.126		**0**	**0.974**	0.974	
Durables	0.6711	no	0.013	0.9215	0.952	0.884		**0**	**0.99**	0.996	
MKT	0	no	-0.152	0.8754	0.9273	0.592		-1.03	0.001	0.359	
24 portfolios sorted by BM within industry											
Intercept	-	-	1.767	0.0404	0.061	0.414	11	1.317	0	0.344	3
Nondur	0.1513	no	0.232	0.029	0.0579	0.526		**0**	**0.993**	0.995	
Durables	0.6878	no	-0.002	0.9891	0.9902	0.738		**0**	**0.976**	0.998	
MKT	0	no	0.44	0.5977	0.6831	0.46		0.89	0.002	0.488	
25 portfolios sorted by MKT and HML betas											
Intercept	-	-	1.558**	0.0231	0.1066	0.014	44	2.185***	0	0.004	1
Nondur	0.9222	no	0.522	0.0009	0.0206	0.272		**0**	**0.999**	0.999	
Durables	0.021	no	0.112	0.3823	0.5434	0.456		**0**	**0.996**	0.998	
MKT	0	no	0.338	0.6122	0.7587	0.842		-0.169	0	0.942	

The case of nontradables

Equity premium puzzle: a long-standing problem of low correlation between consumption growth and financial markets, e.g. Mehra and Prescott (1985).

Investment factors, human capital proxies, *cay*, broker-dealer leverage, Q1-Q4 consumption growth, innovations in volatility, etc...

- Measurement error in the nontradable factors causes attenuation bias in the estimates of the factor exposures (β)
- In finite samples it lowers their size and spread
- The problem seems to be particularly severe for consumption factors
- Large measurement error in data + model misspecificaion call for particular caution

- Cannot be the full story: mimicking portfolios are still weak!

Cross-section of stocks and mimicking portfolios

Cross-section of stocks and non-tradable factors: Yogo (2006) and mimicking portfolios.

			Fama-MacBeth estimator					Pen-FM estimator			
Factors	p-value (Wald)	t_m surv	λ_j	p-value (OLS)	p-value (Shanken)	p-value (Bootstrap)	R^2 (%)	λ_j	Shrinkage rate (Bootstrap)	p-value (Bootstrap)	R^2 (%)
					25 Fama-French portfolios						
Intercept	-	-	2.333**	0.0124	0.0321	0.025	55	3.658***	0	0	21
Nondur	0	0	0.136	0.0096	0.0346	0.078		**0**	**0.968**	0.968	
Durables	0	0	-0.019	0.5011	0.6161	0.96		**0**	**0.9995**	0.9995	
MKT	0	0	-0.19	0.8433	0.88	0.722		-1.252	0	0.217	
					24 portfolios sorted by BM within industry						
Intercept	-	-	1.768	0.0403	0.0453	0.256	11	2.249*	0	0.096	0
Nondur	0	no	0.061	0.0101	0.0554	0.208		**0**	**0.965**	0.966	
Durables	0	no	-0.005	0.8829	0.8971	0.923		**0**	**0.9095**	0.9805	
MKT	0	no	0.426	0.6066	0.6773	0.527		-0.039	0	0.819	
					25 portfolios sorted by MKT and HML betas						
Intercept	-	-	1.556**	0.0233	0.0364	0.025	44	2.285***	0	0.002	4
Nondur	0	no	0.098*	0.0003	0.0046	0.053		**0**	**0.989**	0.989	
Durables	0	no	0.021	0.4692	0.569	0.606		**0**	**0.999**	0.999	
MKT	0	no	0.354	0.595	0.7013	0.752		-0.272	0.0005	0.7405	

Cross-section of stocks and tradable factors

Cross-section of stocks and tradable factors: q-factor model of Hou, Xue and Zhang (2014).

Factors	p-value (Wald)	t_m surv	Fama-MacBeth estimator λ_j	p-value (OLS)	p-value (Shanken)	p-value (Bootstrap)	R^2 (%)	Pen-FM estimator λ_j	Shrinkage rate (Bootstrap)	p-value (Bootstrap)	R^2 (%)
25 portfolios, sorted by size and book-to-market											
Intercept	-	-	1.045***	0.001	0.0018	0.004	77	1.034***	0	0	70
MKT	0	no	-0.553	0.0807	0.16	0.166		-0.505	0	0.184	
M/E	0	yes	0.363**	0	0.0165	0.05		0.255	0.002	0.158	
I/A	0	yes	0.407***	0	0.0022	0.004		0.363**	0.004	0.012	
ROE	0	no	0.494**	0.0148	0.0432	0.042		**0**	**0.822**	0.822	
25 portfolios sorted by value and momentum											
Intercept	-	-	0.256	0.6115	0.6339	0.66	88	0.454	0	0.218	88
MKT	0	no	0.285	0.5489	0.6024	0.604		0.105	0.001	0.921	
M/E	0	yes	0.5***	0	0.0014	0.004		0.482***	0	0.006	
I/A	0	no	0.063	0.796	0.8174	0.788		**0**	**0.759**	0.979	
ROE	0	yes	0.665***	0	0.0006	0.006		0.63***	0	0.004	
10 portfolios sorted on momentum											
Intercept	-	-	1.164	0.1222	0.1502	0.432	93	-0.064	0	0.582	90
MKT	0	no	-0.631	0.3834	0.4327	0.73		0.578	0.001	0.951	
M/E	0	yes	0.73	0.3213	0.3632	0.614		**0**	**0.968**	0.968	
I/A	0	no	0.02	0.9685	0.971	0.91		**0**	**0.582**	0.6	
ROE	0	yes	0.468	0.1425	0.1961	0.206		0.742**	0.005	0.033	
19 portfolios sorted by P/E ratio											
Intercept	-	-	2.71	0.0611	0.1233	0.504	81	0.2578	0	0.544	76
MKT	0	yes	-2.124	0.1293	0.2153	0.7		0.272	0	0.968	
M/E	0	yes	1.132	0.0447	0.1056	0.54		**0**	**0.957**	0.967	
I/A	0	no	0.056	0.8144	0.8527	0.374		0.443*	0.051	0.095	
ROE	0	no	0.072	0.798	0.842	0.946		**0**	**0.669**	0.845	

Conclusion: Asset pricing with spurious factors

Spurious factors can be a pervasive problem for linear factor models, as data mining.

Availability of new data and better computational methods does not make it easier, it makes it worse!

- Many linear factor models seem to be weakly identified:
  - consumption (in particular, durables and consumption volatility), labour, cay
  - some currency factors
- Measurement error in nontradables cannot be fully responsible for this result

How widespread is the problem empirically? What about the nonlinear models?

Should we use individual stocks instead of portfolios?

Concern 1: What about implied factors?

- Fundamental theorem of asset pricing implies that there is a factor that explains cross-sectional differences in asset returns:

$$\mathbb{E}\left[R^e_{t+1}\right] = -\frac{cov(R^e_{t+1}, M_{t,t+1})}{\mathbb{E}\left[M^e_{t,t+1}\right]}$$

- Why not extract common factors directly from the cross-section of returns?

Concern 1: What about implied factors?

- Fundamental theorem of asset pricing implies that there is a factor that explains cross-sectional differences in asset returns:

$$\mathbb{E}\left[R^e_{t+1}\right] = -\frac{cov(R^e_{t+1}, M_{t,t+1})}{\mathbb{E}\left[M^e_{t,t+1}\right]}$$

- Why not extract common factors directly from the cross-section of returns?
- PCA, ICA, etc focus on the time series dimension of returns and cross-correlations:
  - Nothing in the SDF representation implies that the **only** source of correlation between returns is due to their loadings on the pricing kernel
  - Nothing in the SDF representation implies that asset returns cannot load on other factors that come with zero price of risk
  - Nothing in the SDF representation implies normality or linearity of SDF in terms of the asset returns
- Empirically, PCA and related factor extraction techniques focus on the time series dynamics only, and do a really bad job at explaining the cross-section of expected returns
- Fundamental factors, in turn, are successful at capturing the cross-sectional aspect, but usually lose out to PCA in terms of the time-series R^2.
- Tradeoff between TS and CS asset pricing.

Concern 2: Why not use all the individual stocks?

- Historically done due to computational burden
- Forming portfolios loses out on the useful information contained in the cross-section, but allows to diminish idiosyncratic noise
- Individual stock returns are very idiosyncratic/noisy, it is hard to identify systematic sources of risk
- Typically a custom cross-section is created for a new factor, to highlight time series exposure to it.

Gagliardini, Scallet and Ossola (2016): linear factor models on a large cross-section of stocks (large N, large T asymptotics), very flexible approach:

- Fama-MacBeth regressions
- Unbalanced panel
- Time variation in betas
- Time variation in risk premia

Reduced to structural

- Economists study human behaviour: people react to external shocks and change their response
- Careful interpretation of the reduced form findings: it may not be easy all the general equilibrium effects
- Example: skill of the mutual funds managers, Berk and van Binsbergen (2014)
  - Funds have to report their quarterly holdings (13F form)

Reduced to structural

- Economists study human behaviour: people react to external shocks and change their response
- Careful interpretation of the reduced form findings: it may not be easy all the general equilibrium effects
- Example: skill of the mutual funds managers, Berk and van Binsbergen (2014)
    - Funds have to report their quarterly holdings (13F form)
    - 200 GB of data!
    - Vast empirical evidence: manager's skill (α) does not seem to be persistent over years
    - Does that mean all those returns are simply due to luck?

Reduced to structural

- Economists study human behaviour: people react to external shocks and change their response
- Careful interpretation of the reduced form findings: it may not be easy all the general equilibrium effects
- Example: skill of the mutual funds managers, Berk and van Binsbergen (2014)
  - Funds have to report their quarterly holdings (13F form)
  - 200 GB of data!
  - Vast empirical evidence: manager's skill (α) does not seem to be persistent over years
  - Does that mean all those returns are simply due to luck?
  - no, skill is persistent, but markets are fast to catch up and there is decreasing returns to scale
  - Consider a manager who could generate \$1 mln profit on a portfolio of \$10 mln ($\alpha = 10\%$).
  - Next year he has to manage the portfolio of \$20 mln, and the same \$1 mln profit is only 5%.
  - Similar to 'hot hand fallacy' in basketball

Spurious or traded away?

Does academic research have a direct impact on the financial industry?

Spurious or traded away?

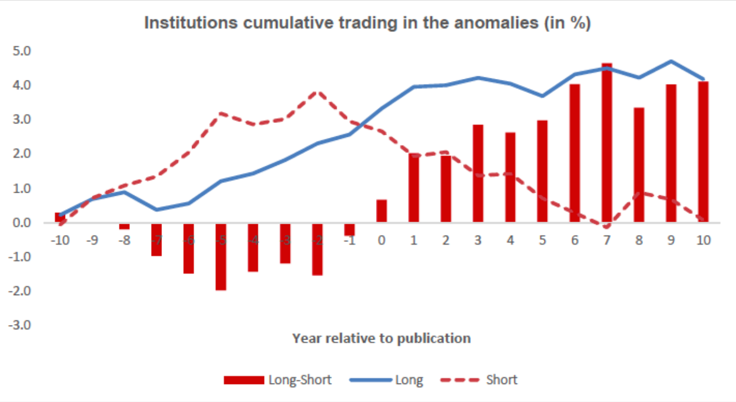
Does academic research have a direct impact on the financial industry?

- McLean and Pontiff (2016): returns to the trading strategies post publication are over 50% lower
- Two potential reasons: sample selection effect and arbitrageurs activity.
- Calluzzo, Moneta and Topaloglu (2016) study institutional trading around the publication of anomalies.

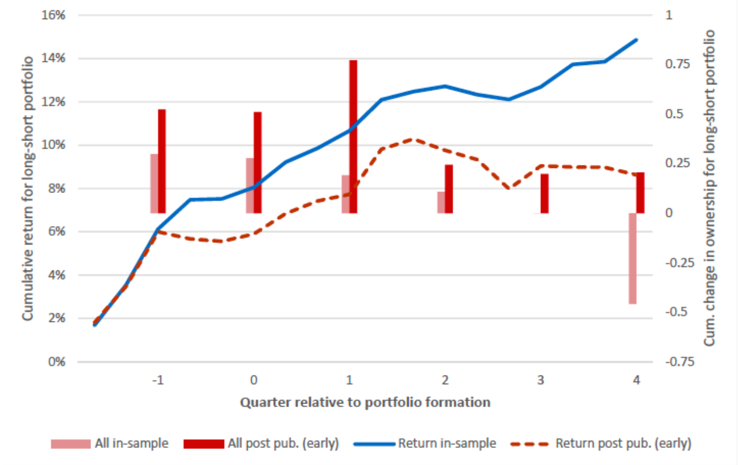
Main findings:

- Focus on the long and short portfolios, corresponding to 14 prominent asset pricing anomalies.
- For the annual anomalies, 0.75% of the total net ownership in the long/short portfolio ($8.55 bln change in ownership).
- The increase is primarily driven by hedge funds and transient institutions.
- *Ex ante* portfolio returns are substantially larger than those of the *Ex post* strategy (when the anomaly is publicly known).

Trading patterns



Returns patterns



Conclusion

- Availability of new data and better computational power (optimization, scraping, etc) led to a tremendous growth in data-driven research
- Finance was a particular object of interest:
  - factor-based trading
  - algorithmic trading
  - high-frequency data
  - new datasets
- It is easy to find many spurious relationships, but miss out on the important things
  - Get your econometrics right!
  - Think of the interpretation of the findings.