

Empirical Asset Pricing

Risk factors out-of-sample

Jonas Nygaard Eriksen

Department of Economics and Business Economics
Aarhus University and CREATES

Spring 2022

- ➊ Introduction
- ➋ McLean and Pontiff (2016)
- ➌ Linnainmaa and Roberts (2018)
- ➍ Other perspectives in the debate
- ➎ Potential projects

- * There is **great debate** about the **number and identification** of reliable **risk factors** and **cross-sectional return predictors** (Harvey et al., 2016, Harvey, 2017, McLean and Pontiff, 2016, Linnainmaa and Roberts, 2018, Engelberg et al., 2018, Calluzzo et al., 2019, Chen and Zimmermann, 2020, Chen, 2021a, Chen and Zimmermann, 2021, Jensen et al., 2021, Harvey and Liu, 2021)

The opposing sides of the debate

Dark side: Factors are false

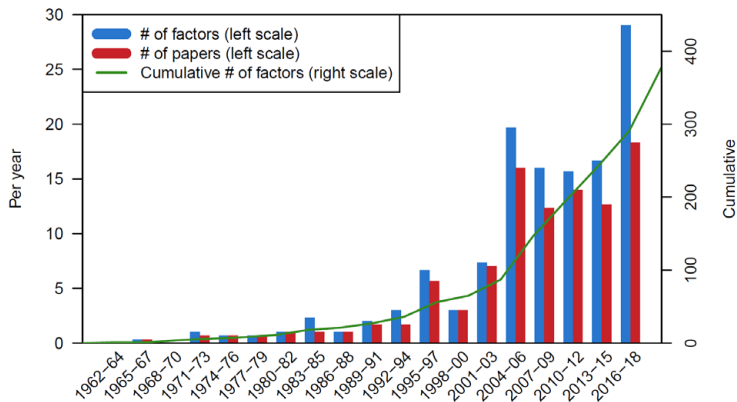
1. Data-mining
2. *p*-hacking
3. Chance result
4. Replication crisis

Bright side: Factors are real

1. Economically motivated
2. Robustness checks
3. Mispricing
4. Factors replicate

Factor production is rampant

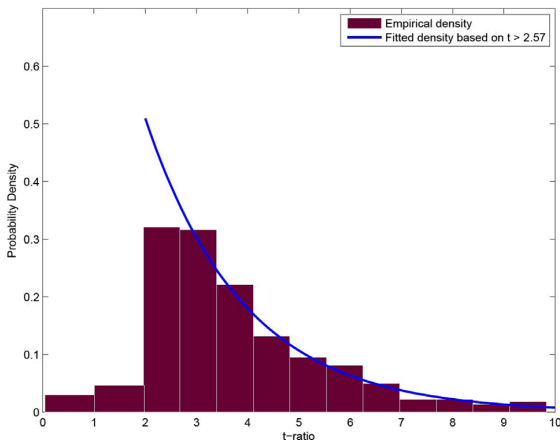
- * Harvey and Liu (2019) provide a [census](#) of the [proverbial factor zoo](#) (Cochrane, 2011, Harvey et al., 2016) [Google sheet with factors](#)
- * The below figure details the [factor production](#) from 1963–2018 and clearly indicates a strong increase (382 factors in top journals alone)



* Journals published through December 2018. Data collection in January 2019.

Distribution of reported t -statistics

- * Harvey (2017) argues that the **distribution of reported t -statistics** is indicative of a **publication bias** and **p -hacking**
- * In particular, note that t -statistics in the range of 2.0 to 2.57 is almost the same as the number reporting in the range of 2.57 to 3.14. Also, notice that very few papers with negative results (t -statistic less than 2.00) are published



- * McLean and Pontiff (2016) ask the following question: Does cross-sectional return predictability typically persist post-publication?
- * To investigate this, they consider 97 cross-sectional return predictors (or risk factors) over three distinct periods
 1. The original study's sample period
 2. The out-of-sample period (post-sample, pre-publication)
 3. The post-publication period
- * Examining the behavior of returns over these three periods, we can distinguish between three possible explanations
 1. Statistical bias \Rightarrow Predictability should disappear out-of-sample
 2. Rational risk \Rightarrow Predictability should be the same in-sample, out-of-sample, and post-publication
 3. Mispricing \Rightarrow If arbitrage is costless, then it should disappear completely. If arbitrage is costly, then the effect should at least decay post-publication

Summary of main findings

- * The **main findings** in McLean and Pontiff (2016) can be **summarized** as follows
 1. The average predictor's long-short return **declines** by **26% out-of-sample**
 - ┆ One can view this as an upper bound on the effect of statistical biases!
 2. The average predictor's long-short return **shrinks 58% post-publication**
 - ┆ Together with the 26% out-of-sample decline, this implies a lower bound on the publication effect of about 32%
 - ┆ Rejection of the hypothesis that return predictability disappears completely and that predictability post-publication is unchanged
 3. The **decay in portfolio returns** is larger for predictor portfolios with **higher in-sample returns and t -statistics**
 4. Post-publication returns are **lower** for predictors that are **less costly to arbitrage**, i.e., portfolios concentrated in liquid and low idiosyncratic risk stocks
 5. Publication affects the **correlation** between **predictor portfolio returns**. (Yet-to-be) published predictors are highly correlated with (yet-to-be) published
- ⇒ The empirical evidence suggests that investors **learn about mispricing** from **academic publications**

Summary statistics

- * The below table provides an [overview of the characteristics](#) of the data. Note that **only 85 of the 97 predictors** have a t -statistic above 1.5

Number of predictor portfolios	97
Predictors portfolios with t -statistic > 1.5	85 (88%)
Mean publication year	2000
Median publication year	2001
Predictors from finance journals	68 (70%)
Predictors from accounting journals	27 (28%)
Predictors from economics journals	2 (2%)
Mean portfolio return in-sample	0.582
Standard deviation of mean in-sample portfolio return	0.395
Mean observations in-sample	323
Mean portfolio return out-of sample	0.402
Standard deviation of mean out-of-sample portfolio return	0.651
Mean observations out-of-sample	56
Mean portfolio return post-publication	0.264
Standard deviation of mean post-publication portfolio return	0.516
Mean observations post-publication	156

Cross-sectional predictors

- * In their **Internet Appendix**, they provide a full list of **predictors and their classification** into four broad categories (plus definitions)

Event	Market	Valuation	Fundamental
Change in Asset Turnover	52-Week High	Advertising/MV	Accruals
Change in Profit Margin	Age-Momentum	Analyst Value	Age
Change in Recommendation	Amihud's Measure	Book-to-Market	Asset Growth
Chg. Forecast + Accrual	Beta	Cash Flow/MV	Asset Turnover
Debt Issuance	Bid/Ask Spread	Dividends	Cash Flow Variance
Dividend Initiation	Coskewness	Earnings-to-Price	Earnings Consistency
Dividend Omission	Idiosyncratic Risk	Enterprise Component of B/P	Forecast Dispersion
Dividends	Industry Momentum	Enterprise Multiple	G Index
Down Forecast	Lagged Momentum	Leverage Component of B/P	Gross Profitability
Exchange Switch	Long-term Reversal	Marketing/MV	G-Score
Growth in Inventory	Max	Org. Capital	G-Score 2
Growth in LTNOA	Momentum	R&D/MV	Herfindahl
IPO	Momentum and Long-term Reversal	Sales/Price	Investment
IPO + Age	Momentum-Ratings		Leverage
IPO no R&D	Momentum-Reversal		M/B and Accruals
Mergers	Price		NOA
Post Earnings Drift	Seasonality		Operating Leverage
R&D Increases	Short Interest		O-Score
Ratings Downgrades	Short-term Reversal		Pension Funding
Repurchases	Size		Percent Operating Accrual
Revenue Surprises	Volume		Percent Total Accrual
SEO	Volume Trend		Profit Margin
Share Issuance 1-Year	Volume Variance		Profitability
Share Issuance 5-Year	Volume-Momentum		ROE
Spinoffs	Volume/MV		Sales Growth
Sustainable Growth			Tax
Total External Finance			Z-Score
Up Forecast			
$\Delta \text{Capex} - \Delta \text{Industry CAPEX}$			
$\Delta \text{Noncurrent Op. Assets}$			
$\Delta \text{Sales} - \Delta \text{Inventory}$			
$\Delta \text{Sales} - \Delta \text{SG\&A}$			
$\Delta \text{Work. Capital}$			

Returns post-sample and post-publication

- * The **main panel regression** of the paper is specified as follows (where α_i captures **predictor fixed effects**)

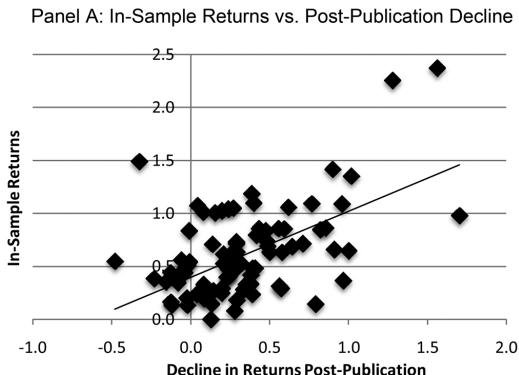
$$R_{it} = \alpha_i + \beta_1 \text{Post sample dummy}_{it} + \beta_2 \text{Post publication dummy}_{it} + e_{it} \quad (1)$$

Variables	(1)	(2)	(3)	(4)
Post-Sample (S)	-0.150*** (0.077)	-0.180** (0.085)	0.157 (0.103)	0.067 (0.112)
Post-Publication (P)	-0.337*** (0.090)	-0.387*** (0.097)	-0.002 (0.078)	-0.120 (0.114)
S × Mean			-0.532*** (0.221)	
P × Mean			-0.548*** (0.178)	
S × <i>t</i> -statistic				-0.061*** (0.023)
P × <i>t</i> -statistic				-0.063*** (0.018)
Predictor FE?	Yes	Yes	Yes	Yes
Observations	51,851	45,465	51,851	51,944
Predictors (<i>N</i>)	97	85	97	97
Null : S = P	0.024	0.021		
Null: P = -1 × (mean)	0.000	0.000		
Null: S = -1 × (mean)	0.000	0.000		

- * We have **reductions in returns** of $15/58.2 = 26\%$ and $33.7/58.2 = 58\%$, respectively, relative to the **average in-sample return of 58.2bps**

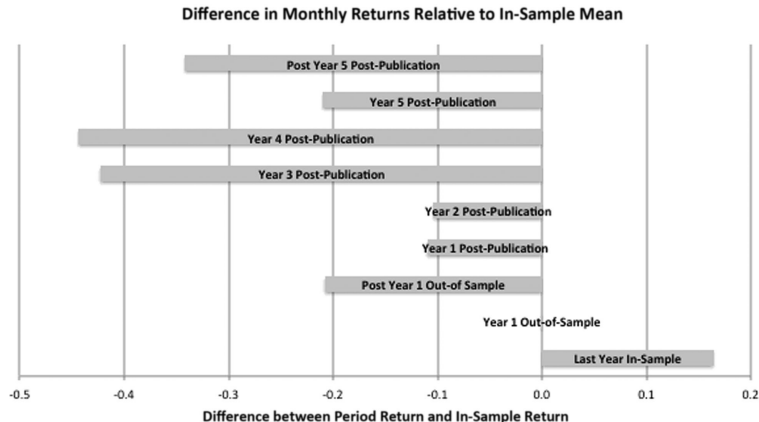
In-sample returns and post-publication decline

- * Higher in-sample returns are correlated with larger reductions in post-publication returns (similar for t -statistics)



Returns around sample-end and publication dates

- * We then examine **changes in predictability** by considering a set of **partitions of the sample in event time** at higher granularity



Time trends and persistence in returns

- * Are the results due to **time trends**, **persistencies**, or **lower costs** of corrective trading instead of **academic research**?
- * The results are not supportive of this hypothesis

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Time	-0.069*** (0.011)		-0.069*** (0.026)			
Post-1993		-0.120 (0.074)	0.303*** (0.118)			
Post-Sample			-0.190** (0.081)	-0.179** (0.080)	-0.132* (0.076)	-0.128 (0.078)
Post-Publication			-0.362*** (0.124)	-0.310** (0.122)	-0.295*** (0.089)	-0.258*** (0.093)
1-Month Return					0.114*** (0.015)	
12-Month Return						0.020*** (0.004)
Observations	51,851	51,851	51,851	51,851	51,754	50,687
Char. FE?	Yes	Yes	Yes	Yes	Yes	Yes
Time FE?	No	No	No	Yes	No	No

Returns across predictor types

- * To test for **differences among predictor groups**, we can run the following **regression model**

$$R_{it} = \alpha_i + \beta_1 \text{Post publication dummy}_{it} + \beta_2 \text{Predictor type dummy}_i + \beta_3 \text{Post publication dummy}_{it} \times \text{Predictor type dummy}_i + e_{it} \quad (2)$$

Variable	(1)	(2)	(3)	(4)
Post-Publication (P)	-0.208*** (0.059)	-0.316*** (0.097)	-0.310*** (0.080)	-0.301*** (0.089)
Market	0.304*** (0.079)			
P × Market	-0.244 (0.169)			
Event		-0.098** (0.046)		
P × Event		0.105 (0.091)		
Valuation			-0.056 (0.063)	
P × Valuation			0.186 (0.131)	
Fundamental				-0.201*** (0.045)
P × Fundamental				0.025 (0.089)
Constant	0.482*** (0.036)	0.606*** (0.052)	0.585*** (0.000)	0.630*** (0.053)
Observations	51,851	51,851	51,851	51,851
Predictors	97	97	97	97
Type + (P × Type)	0.060	0.007	0.121	-0.176
p-value	0.210	0.922	0.256	0.012

- * The results so far are consistent with the idea that **publication attracts arbitrageurs**, which leads to **lower returns post-publication**
- * **Costly arbitrage** can **prevent mispricing from being fully eroded**, which suggests that predictor portfolios concentrated in **stocks that are costlier to arbitrage** (e.g., smaller stocks, illiquid stocks, and high idiosyncratic risk stocks) **should decline less post-publication**
- * McLean and Pontiff (2016) consider **three transaction costs variables** (size, bid-ask spreads, dollar volume) and **two holding cost variables** (idiosyncratic risk, dividend-payer dummy) and run the following regression model

$$R_{it} = \alpha_i + \beta_1 \text{Post publication dummy}_{it} + \beta_2 \text{Arbitrage cost}_i + \beta_3 \text{Post publication dummy}_{it} \times \text{Arbitrage cost}_i + e_{it} \quad (3)$$

Costly arbitrage results

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Post-Pub. (P)	-0.190 (0.274)	-0.139 (0.235)	0.215 (0.230)	-0.242 (0.273)	-0.321 (0.211)	-0.264** (0.078)
P × Size	-0.138 (0.459)					
Size	-1.064** (0.236)					
P × Spreads		-0.301 (0.603)				
Spreads		1.228** (0.252)				
P × Dol.Vol.			-1.059* (0.500)			
Dol. Vol.			0.215 (0.308)			
P × Idio. Risk				-0.047 (0.554)		
Idio. Risk				2.064*** (0.330)		
P × Div.					-0.321 (0.211)	
Div.					-0.526*** (0.145)	
P × Index						-0.009 (0.019)
Index						-0.056*** (0.011)
Constant	1.145*** (0.130)	0.146* (0.174)	0.476*** (0.144)	-0.469*** (0.171)	0.855*** (0.097)	0.565*** (0.000)
Observations	51,851	51,851	51,851	51,851	51,851	51,851
CA + (P × CA)	-1.202	0.927	-0.844	2.017	-0.847	-0.065
p-value	0.003	0.096	0.000	0.000	0.144	0.000

Post-sample and -publication trading activity dynamics

- * If **academic publication** provides market participants with information, then **informed trading activity** should affect indicators of trading beyond prices
- * Below, we investigate whether **variance, trading volume, dollar trading volume, and short interest** increase in predictor portfolios post-publication
- * Note that a **significant post-sample coefficient** is indicative of the **informational content being out prior** to publication (e.g., working papers and conferences)

Variables	Variance	Trading volume	Dollar volume	Short-long short interest
Post-Sample (S)	-0.054*** (0.007)	0.092*** (0.001)	0.066*** (0.007)	0.166*** (0.014)
Post-Publication (P)	-0.065*** (0.008)	0.187*** (0.013)	0.097*** (0.007)	0.315*** (0.013)
Observations	52,632	52,632	52,632	41,026
Time FE?	Yes	Yes	Yes	No
Predictor FE?	Yes	Yes	Yes	Yes
Null: S = P	0.156	0.000	0.000	0.000

Correlation structure and publication

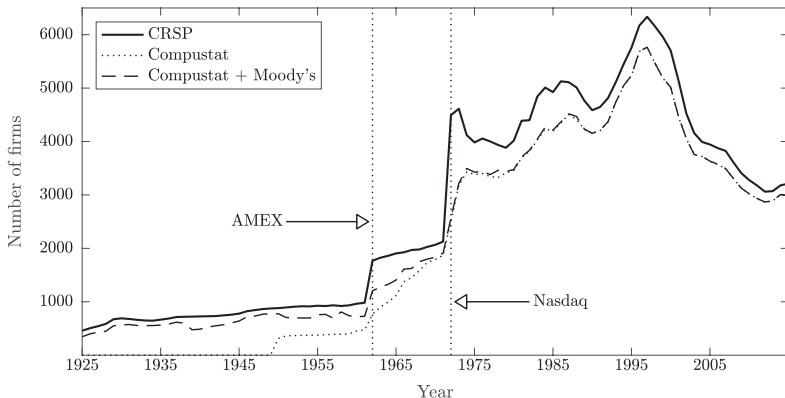
- * If predictor returns **reflect mispricing**, and if **mispricing has a common source**, then we should expect that
 1. In-sample returns are significantly related to other in-sample portfolios
 2. Once a predictor is published (and attracts arbitrageurs), its return should correlate (less) more with the returns of other (pre-) post-publication predictors
- * We examine this by **regressing predictor portfolio returns** on the returns of an
 1. equal-weighted index of all other portfolios that are pre-publication
 2. equal-weighted index of all of the other portfolios that are post-publication
 3. A post-publication dummy and interactions with the indices

Variables	Coefficients
In-Sample Index Returns	0.748*** (0.027)
Post-Publication Index Return	-0.008 (0.004)
$P \times \text{In-Sample Index Returns}$	-0.674*** (0.033)
$P \times \text{Post-Publication Index Return}$	0.652*** (0.045)
Publication (P)	-0.081 (0.042)
Constant	0.144*** (0.019)
Observations	42,975
Predictors	97

- * Linnainmaa and Roberts (2018) ask the following question: Does accounting-based anomalies persist out-of-sample (both pre- and post-sample)?
- * To examine this, they consider a broad selection of 36 accounting-based anomalies over three distinct periods
 1. “In-sample” denotes the sample frame used in the original discovery of an anomaly
 2. “Pre-sample” denotes the sample frame occurring prior to the in-sample period
 3. “Post-sample” denotes the sample frame occurring after the in-sample period
- * Similar to McLean and Pontiff (2016), they distinguish between three possible explanations
 1. Unmodeled risk \Rightarrow Stocks risk are multidimensional and models are misspecified
 2. Mispricing \Rightarrow Investor irrationality and limits to arbitrage cause anomaly returns
 3. Data-snooping \Rightarrow Anomalies are artifacts of chance error
- * Main conclusion: Most anomaly returns are decidedly an in-sample phenomenon consistent with data-snooping (i.e., most risk factors are false)

Number of firms available

- * One of the main contribution of Linnainmaa and Roberts (2018) is to **back-fill accounting information** using **Moody's Industrial and Railroad manuals** to obtain out-of-sample data points back in time
- * See also Baltussen et al. (2021) for a paper constructing a pre-CRSP dataset



Case study: Fama-French five factors

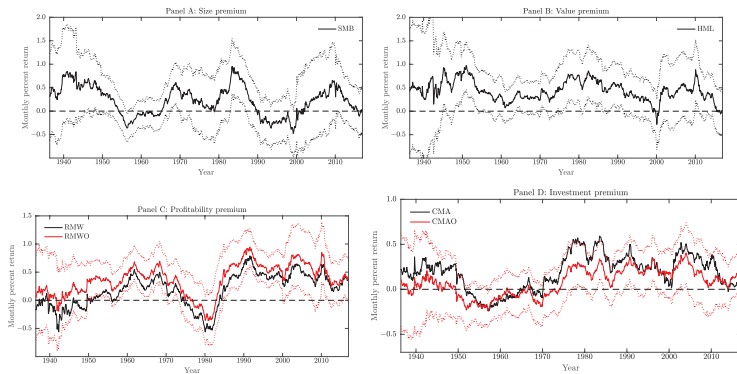
- * To **illustrate their approach**, consider the case of the Fama and French (2015) five-factor model using **pre-1963 accounting data**

Panel A: Monthly percent returns

Portfolio	Pre-1963 sample			
	1926:7 –1938:6	1938:7 –1963:6	1926:7 –1963:6	1963:7 –2016:12
Portfolios sorted by size and book-to-market				
Small Growth	0.73	1.17	1.03	0.88
Neutral	1.04	1.32	1.23	1.29
Value	1.41	1.62	1.55	1.40
Big Growth	0.80	0.99	0.93	0.88
Neutral	0.73	1.16	1.02	0.96
Value	0.88	1.51	1.31	1.13
Size factor	0.26 (0.62)	0.15 (1.05)	0.18 (1.11)	0.20 (1.54)
Value factor	0.38 (0.66)	0.48 (2.87)	0.45 (2.06)	0.38 (3.40)
Portfolios sorted by size and profitability				
Small Weak	0.88	1.32	1.18	0.96
Neutral	0.78	1.38	1.19	1.24
Robust	0.75	1.37	1.17	1.31
Big Weak	0.81	1.09	1.00	0.83
Neutral	0.76	1.10	0.99	0.90
Robust	0.68	1.23	1.05	1.03
Profitability factor	–0.13 (–0.31)	0.09 (0.64)	0.02 (0.14)	0.28 (3.09)
Portfolios sorted by size and investment				
Small Aggressive	0.84	1.23	1.10	0.98
Neutral	1.58	1.36	1.43	1.34
Conservative	1.06	1.36	1.27	1.33
Big Aggressive	0.69	1.11	0.97	0.90
Neutral	1.14	1.12	1.13	0.94
Conservative	0.86	1.04	0.98	1.08
Investment factor	0.19 (0.79)	0.03 (0.32)	0.09 (0.80)	0.26 (3.28)

The dynamics of FF5 factor premiums

- * Consider the **ten-year rolling window** average returns to the Fama and French (2015) **factors**
- * See also Fama and French (2021) for a discussion of the recent performance of the value premium



Risk factors under consideration

- * The **36 accounting-based anomalies** under consideration, and **their groupings**, are specified below and information about the construction can be found in the paper appendix

Category	No.	Anomaly	Original study	Original sample
Profitability	1	Gross profitability	Novy-Marx (2013)	1963–2010
	2	Operating profitability*	Fama and French (2015)	1963–2013
	3	Return on assets*	Haugen and Baker (1996)	1979–1993
	4	Return on equity*	Haugen and Baker (1996)	1979–1993
	5	Profit margin	Soliman (2008)	1984–2002
Earnings quality	6	Change in asset turnover	Soliman (2008)	1984–2002
	7	Accruals*	Sloan (1996)	1962–1991
	8	Net operating assets	Hirshleifer, Hou, Teoh, and Zhang (2004)	1964–2002
	9	Net working capital changes	Soliman (2008)	1984–2002
Valuation	10	Book-to-market	Fama and French (1992)	1963–1990
	11	Cash flow / price	Lakonishok, Shleifer, and Vishny (1994)	1968–1990
	12	Earnings / price	Basu (1977)	1957–1971
	13	Enterprise multiple*	Loughran and Wellman (2011)	1963–2009
	14	Sales / price	Barbee, Mukherji, and Raines (1996)	1979–1991
Investment and growth	15	Asset growth	Cooper, Gulen, and Schill (2008)	1968–2003
	16	Growth in inventory	Thomas and Zhang (2002)	1970–1997
	17	Sales growth	Lakonishok, Shleifer, and Vishny (1994)	1968–1990
	18	Sustainable growth	Lockwood and Prombutr (2010)	1964–2007
	19	Adjusted CAPX growth*	Abarbanell and Bushee (1998)	1974–1993
	20	Growth in sales – inventory	Abarbanell and Bushee (1998)	1974–1993
	21	Investment growth rate*	Xing (2008)	1964–2003
	22	Abnormal capital investment*	Titman, Wei, and Xie (2004)	1973–1996
	23	Investment to capital*	Xing (2008)	1964–2003
	24	Investment-to-assets	Lyandres, Sun, and Zhang (2008)	1970–2005
Financing	25	Debt issuance*	Spiess and Affleck-Graves (1999)	1975–1994
	26	Leverage	Bhandari (1988)	1948–1979
	27	One-year share issuance	Pontiff and Woodgate (2008)	1970–2003
	28	Five-year share issuance	Daniel and Titman (2006)	1968–2003
	29	Total external financing*	Bradshaw, Richardson, and Sloan (2006)	1971–2000
Distress	30	O-score	Dichev (1998)	1981–1995
	31	z-score*	Dichev (1998)	1981–1995
	32	Distress risk	Campbell, Hilscher, and Szilagyi (2008)	1963–2003
Other	33	Industry concentration	Hou and Robinson (2006)	1951–2001
Composite anomalies	34	Piotroski's F-score	Piotroski (2000)	1976–1996
	35	M/B and accruals*	Bartov and Kim (2004)	1981–2000
	36	QMJ: Profitability	Asness, Frazzini, and Pedersen (2013)	1956–2012

Returns in pre-, in-, and post-sample periods

- * The **main result** of Linnainmaa and Roberts (2018) is that **anomaly returns are a decidedly in-sample phenomenon**, which is clearly evident in their Table 6

Measure	Pre-sample	In-sample	Post-sample	Differences		
				Pre – In	Post – In	Post – Pre
<i>Panel A: Full pre-1963 sample</i>						
Average returns						
Average return	0.08 (2.21)	0.29 (7.01)	0.09 (1.72)	−0.21 (−3.78)	−0.20 (−3.69)	0.00 (0.03)
Sharpe ratio	0.15 (3.38)	0.54 (7.57)	0.13 (1.52)	−0.39 (−4.71)	−0.42 (−4.14)	−0.03 (−0.30)
CAPM						
Alpha	0.15 (4.80)	0.34 (9.75)	0.17 (3.50)	−0.20 (−4.27)	−0.18 (−3.44)	0.02 (0.38)
Information ratio	0.22 (5.08)	0.66 (9.72)	0.27 (2.99)	−0.43 (−5.43)	−0.40 (−3.83)	0.04 (0.43)
Three-factor model						
Alpha	0.17 (6.42)	0.27 (10.12)	0.12 (3.19)	−0.10 (−2.57)	−0.15 (−3.44)	−0.05 (−1.10)
Information ratio	0.28 (6.35)	0.60 (9.91)	0.25 (2.86)	−0.32 (−4.26)	−0.35 (−3.46)	−0.03 (−0.32)

Effect of state date on returns

- * To investigate the **sensitivity to the in-sample period**, Linnainmaa and Roberts (2018) run a panel regression of the form

$$\text{anomaly}_{it} = \beta_0 + \beta_1 \text{Pre-sample}_{it} + \mu_i + \varepsilon_{it} \quad (4)$$

where μ_i is an **anomaly fixed effect** and Pre-sample_{it} is an **indicator equal to one** if the anomaly-month observation falls in the time period before the start date of the anomaly's in-sample start date

Start year	Average return		CAPM alpha		FF3 alpha		No. of obs.
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\beta}_1$	
1963	0.30 (6.77)	-0.15 (-2.16)	0.36 (10.07)	-0.18 (-2.97)	0.27 (10.35)	-0.14 (-3.18)	14,793
1964	0.30 (6.77)	-0.15 (-1.93)	0.36 (10.06)	-0.19 (-2.86)	0.28 (10.40)	-0.13 (-2.68)	14,385
1965	0.30 (6.78)	-0.13 (-1.58)	0.36 (10.07)	-0.17 (-2.42)	0.28 (10.40)	-0.11 (-2.27)	13,977
1966	0.30 (6.84)	-0.13 (-1.46)	0.37 (10.15)	-0.17 (-2.26)	0.28 (10.45)	-0.12 (-2.21)	13,569
1967	0.31 (6.89)	-0.09 (-0.98)	0.37 (10.21)	-0.13 (-1.84)	0.29 (10.50)	-0.12 (-2.25)	13,161
1968	0.31 (6.90)	-0.07 (-0.80)	0.37 (10.20)	-0.13 (-2.07)	0.29 (10.45)	-0.13 (-2.31)	12,753
1969	0.31 (6.77)	-0.11 (-1.21)	0.38 (10.09)	-0.17 (-2.53)	0.29 (10.44)	-0.18 (-3.04)	12,345
1970	0.30 (6.53)	-0.22 (-2.68)	0.37 (9.91)	-0.24 (-3.37)	0.29 (10.22)	-0.25 (-4.12)	11,937
1971	0.31 (6.63)	-0.22 (-2.65)	0.37 (9.88)	-0.26 (-3.78)	0.29 (10.02)	-0.28 (-4.60)	11,532
1972	0.32 (6.64)	-0.21 (-2.18)	0.38 (9.86)	-0.26 (-3.24)	0.29 (9.86)	-0.31 (-4.52)	11,136
1973	0.31 (6.38)	-0.24 (-2.20)	0.38 (9.70)	-0.28 (-3.31)	0.31 (10.17)	-0.27 (-3.51)	10,740

- * Linnainmaa and Roberts (2018) argue that **data-snooping works through t -values**, so examining **volatility across the three periods** may be worthwhile
- * Consider the panel regression specification for squared demeaned returns

$$(\text{anomaly}_{it} - \bar{r}_{it})^2 = \beta_0 + \beta_1 \text{In-sample}_{it} + \beta_2 \text{Post-sample}_{it} + \mu_{it} + \varepsilon_{it} \quad (5)$$

- * Estimating the model yields the following **parameter estimates**
 - $\beta_1 = -0.61$ with a t -stat of -2.88
 - $\beta_2 = 0.38$ with a t -stat of 0.79

indicating that **volatility is lower during the in-sample period**, which is consistent with data-snooping contaminating the distribution of in-sample returns

Changes in correlation structure of returns

- * Finally, we can investigate **effects upon the correlation structure** as in McLean and Pontiff (2016) by running the regression

$$\text{anomaly}_{it} = a + b_1 \text{in-sample index}_{-i,t} + b_2 \text{post-sample index}_{-i,t} + b_3 \text{post}_{it} + \text{post}_{it} \times (b_4 \text{in-sample index}_{-i,t} + b_5 \text{post-sample index}_{-i,t}) + e_{i,t}$$

Regressor	Coefficient	t-value
Regression 1: In-sample versus post-sample anomalies		
Intercept	0.05	4.54
Main effects		
In-sample index _{-i,t}	0.74	33.98
Post-sample index _{-i,t}	0.08	7.46
Post _{i,t}	-0.06	-2.23
Interactions		
Post _{i,t} × In-sample index _{-i,t}	-0.53	-13.74
Post _{i,t} × Post-sample index _{-i,t}	0.46	11.19
Adjusted R ²		17.9%
N		15,152
Regression 2: In-sample versus pre-sample anomalies		
Intercept	0.07	4.35
Main effects		
In-sample index _{-i,t}	0.74	28.90
Pre-sample index _{-i,t}	0.07	3.42
Pre _{i,t}	-0.04	-2.09
Interactions		
Pre _{i,t} × In-sample index _{-i,t}	-0.69	-22.72
Pre _{i,t} × Pre-sample index _{-i,t}	0.48	13.68
Adjusted R ²		9.3%
N		13,650

Factors are mostly false

- * There is a large camp, mainly headed by Campbell R. Harvey, that claims that most factors are false and that there is a pronounced replication crisis in finance (Harvey et al., 2016, Harvey, 2017, Harvey and Liu, 2019, 2021)
 - Hou et al. (2020) re-evaluate 452 anomalies and find that mitigating for microcaps using NYSE breakpoints and value-weighted returns leads to a failure rate of 65% using a 1.96 cutoff. It increases to a failure rate of 82% using a cutoff of 2.78
 - Tian (2020) runs a data-mining experiment in which she randomly constructs hundreds of three-factor models. She finds that many outperform well known models from the literature, including those with four and five factor. This suggest that the threshold of factor model success needs to be raised
 - Chordia et al. (2020) use information from over 2 million randomly generated trading strategies (using real data and strategies that survive the publication process) to infer the statistical properties of factor strategies. They compute t -statistic threshold that control for multiple hypothesis testing at 3.8 and 3.4 for time-series and cross-sectional regressions, respectively, Failing to account for multiple hypothesis testing leads to a false rejection about 45% of the time

Factors are mostly true

- * There is similarly a growing literature arguing that most factors are true and can be successfully replicated
 - Engelberg et al. (2018) find that **anomaly returns are 50% higher** on **corporate news days** and **six times higher** on **earnings announcement days**. They argue that the results point to the idea that anomaly returns are driven by biased expectations that are (partly) corrected by information arrivals
 - Jacobs and Müller (2020) extend the work of McLean and Pontiff (2016) by **investigating 241 anomalies in 39 countries**. Their results similar point to mispricing rather than data mining
 - Calluzzo et al. (2019) show that there is an **increase in anomaly-related trading** when information about the anomalies is **readily available through academic publications**
 - Chen and Zimmermann (2020) argue that **bias-adjusted returns are only 12.3% smaller** than in-sample returns, which is **well within the** McLean and Pontiff (2016) **upper bounds** and points to mispricing
 - Chen (2021a,b) provides **thought experiments** and alternative views on the results of Harvey et al. (2016) and argues that *p*-hacking cannot be as widely applied as argued and that **most factors are true** as a consequence
 - Jensen et al. (2021) **challenge the dire view** of finance research. They develop and estimate a Bayesian model of factor replication, which leads to different conclusions. They find a **baseline replication rate as high as 56.9%** (plus make available a new global dataset at <https://www.bryankellyacademic.org> under Data)

- * Last, I want to direct attention to a great initiative and a compelling voice in the debate about replicability in finance: <https://www.openassetpricing.com>
- * Chen and Zimmermann (2021) provide open source code, characteristics data, and portfolio returns for about 200 factor portfolios
- * This is a great source of data for re-examining cross-sectional predictors by combining their data with the CRSP database (you can do that using PERMNO)

- * Investigate the returns and pricing ability pre- and post-publication to selected anomalies/factors
- * Investigate whether there are publication effects in other asset classes (e.g., currencies or bonds)
- * Combine a CRSP-based portfolio sorts or a cross-sectional asset pricing exercise with robustness checks inspired by this literature (e.g., re-examine with stricter thresholds)
- * Test whether selected time-series predictors similarly display publication effects inspired by this literature
- * Re-examine selected risk factors/anomalies out-of-sample and investigated whether results can be replicated and extended

References

- Baltussen, G., B. P. van Vliet, and P. van Vliet (2021). The cross-section of stock returns before 1926 (and beyond). Working paper, Erasmus School of Economics.
- Calluzzo, P., F. Moneta, and S. Topaloglu (2019). When anomalies are publicized broadly, do institutions trade accordingly? *Management Science* 65(10), 4451–4949.
- Chen, A. Y. (2021a). The limits of p-hacking: Some thought experiments. *Journal of Finance* 76(5), 2447–2480.
- Chen, A. Y. (2021b). Most claimed statistical findings in cross-sectional return predictability are likely true. Working paper.
- Chen, A. Y. and T. Zimmermann (2020). Publication bias and the cross-section of stock returns. *Review of Asset Pricing Studies* 10(2), 249–289.
- Chen, A. Y. and T. Zimmermann (2021). Open source cross sectional asset pricing. *Critical Finance Review*, Forthcoming.
- Chordia, T., A. Goyal, and A. Saretto (2020). Anomalies and false rejections. *Review of Financial Studies* 33(5), 2134–2179.
- Cochrane, J. H. (2011). Presidential address: Discount rates. *Journal of Finance* 66(5), 1047–1108.
- Engelberg, J., R. D. McLean, and J. Pontiff (2018). Anomalies and news. *Journal of Finance* 73(5), 1971–2001.
- Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics* 116, 1–22.
- Fama, E. F. and K. R. French (2021). The value premium. *Review of Asset Pricing Studies* 11(1), 105–121.
- Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *Journal of Finance* 72(4), 1399–1440.
- Harvey, C. R. and Y. Liu (2019). A census of the factor zoo. Working paper, Duke University.
- Harvey, C. R. and Y. Liu (2021). Uncovering the iceberg from its tip: A model of publication bias and p-hacking. Working paper, Duke University.
- Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies* 29(1), 5–68.
- Hou, K., C. Xue, and L. Zhang (2020). Replicating anomalies. *Review of Financial Studies* 33(5), 2019–2133.
- Jacobs, H. and S. Müller (2020). Anomalies across the globe: One public, no longer existent? *Journal of Financial Economics* 153, 213–230.
- Jensen, T. I., B. Kelly, and L. H. Pedersen (2021). Is there a replication crisis in finance? *Journal of Finance*, Forthcoming.
- Linnainmaa, J. T. and M. R. Roberts (2018). The history of the cross-section of stock returns. *Review of Financial Studies* 31(7), 2606–2649.
- McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *Journal of Finance* 71(1), 5–32.
- Tian, M. (2020). Firm characteristics and empirical factor models: A model mining experiment. *Review of Financial Studies*, Forthcoming.