Does Academic Research Destroy Stock Return Predictability?

David McLean and Jeffrey Pontiff

March 30, 2019

Introduction

Question

What exactly happens to return predictability outside of a study's sample period?

Introduction

What could happen?

1. Perhaps the findings are spurious and only fit the partitioned sample period
2. Or maybe there will be no change, and return predictability will persist
3. Or return predictability will diminish post-publication, suggesting mispricing

Introduction

Hypotheses

1. If findings are spurious, there should be no returns outside of the sample period
2. If return predictability reflects rational expectations, findings should remain
3. If return predictability is due to mispricing, then return predictability should disappear or be decay after a paper is published

Introduction

Key Findings

Using 97 predictors from 79 studies the average predictor long-short quintile portfolio return:

1. Declines 26% out-of-sample
2. Shrinks 58% post-publication

Additionally, decay is larger for:

3. Predictor portfolios with larger in-sample returns and higher t-statistics
4. Predictors constructed from price and trading data

Introduction

Contribution

- Academic publications transmit information to sophisticated investors
- Or *we do something worthwhile and important*!

Research Method

- Exclude time series predictability
- Focus on studies in peer-reviewed finance/economic/accounting literature

Research Method

- Gather 97 cross-sectional relations from 79 studies
- Form long-short quintiles based on results

Some Concerns

- Data no longer available to construct some measures - get proxies
- Rate of increase/decrease - form long-short portfolios with extreme 20 percentiles
- Dummy variables - separate into long or short side of the portfolio
- Discrete values - follow original research

Summary Statistics

Table I

Summary Statistics

This table reports summary statistics for the predictor portfolios studied in this paper. The returns are equal-weighted by predictor portfolio, that is, we first estimate the statistic for each predictor portfolio, and then take an equal-weighted average across the portfolios. The reported standard deviations are the standard deviations of the predictors' mean returns. Our sample period ends in 2013.

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

Empirical Analysis

Baseline Regression:

$R_{it} = \alpha_i + \beta_1$ *Post Sampe Dummy*$_{i,t} + \beta 2$ *Post Publication Dummy*$_{i,t} + e_{it}$

Correlation

- Overall correlation is 0.033, but some will be higher
- Compute standard errors using feasible generalized least squares (FGLS)

Empirical Analysis

Formal Hypotheses

1. If statistical biases are the source of in-sample predictability, then the coefficients on both the post-sample and post-publication dummies should be -0.582
2. If predictors' returns are entirely the result of mispricing and arbitrage resulting from publication corrects all mispricing, the the post-publication coefficient should be equal to -0.582 and the post-sample dummy should not be close to zero
3. If there are no statistical biases and academic papers have no influence on investors' actions, then both of the coefficients should equal zero.

Regression of Predictor Portfolio Returns on Post-Sample and Post-Publication Indicators

Table II

Variables	(1)	(2)
Post-Sample (S)	−0.150***	−0.180**
	(0.077)	(0.085)
Post-Publication (P)	−0.337***	−0.387***
	(0.090)	(0.097)
S × Mean		
P × Mean		
S × *t*-statistic		
P × *t*-statistic		
Predictor FE?	Yes	Yes
Observations	51,851	45,465
Predictors (*N*)	97	85
Null : S = P	0.024	0.021
Null: P = −1 × (mean)	0.000	0.000
Null: S = −1 × (mean)	0.000	0.000

Regression of Predictor Portfolio Returns on Post-Sample and Post-Publication Indicators

Question

Do predictor portfolio returns with higher in-sample means decline more post-publication?

Test

- Interact in-sample means with predictors
- Interaction for post sample is -0.532 and post-sample is 0.157
- Notice $0.157 + (-0.532 \times 0.582) = -0.153$ - nearly the same as column (1)
- Indicates that predictor with larger return could have larger bias
- Or could indicate arbitrageurs are more likely to learn about higher returns before publication

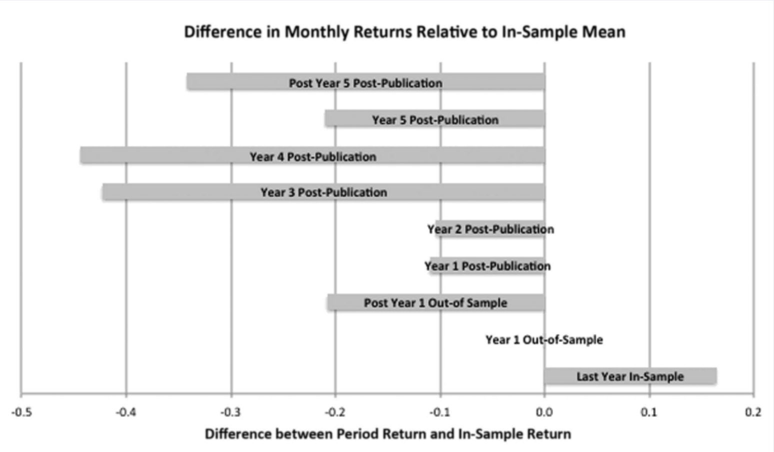
Regression of Predictor Portfolio Returns on Post-Sample and Post-Publication Indicators

Table II

Variables	(1)	(2)	(3)	(4)
Post-Sample (S)	−0.150***	−0.180**	0.157	0.067
	(0.077)	(0.085)	(0.103)	(0.112)
Post-Publication (P)	−0.337***	−0.387***	−0.002	−0.120
	(0.090)	(0.097)	(0.078)	(0.114)
S × Mean			−0.532***	
			(0.221)	
P × Mean			−0.548***	
			(0.178)	
S × *t*-statistic				−0.061***
				(0.023)
P × *t*-statistic				−0.063***
				(0.018)
Predictor FE?	Yes	Yes	Yes	Yes
Observations	51,851	45,465	51,851	51,944
Predictors (*N*)	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		

Result

Figure 2



Controlling for Time Trends and Persistence

Concern

- Perhaps the authors are measuring a time trend that proxies for lower costs of corrective trading

Control

- Use time variable that is equal to 1/100 in January 1926 and increases by 1/100 each month in the sample.

Result

- Time variable has significant and negative slope - suggests portfolio returns have declined overtime.

Controlling for Time Trends and Persistence

Table III

Time Trends and Persistence in Predictor Returns

The regressions reported in this table test for time trends and persistence in predictor returns. *Post-Sample (S)* is equal to one if the month is after the sample period used in the original study and zero otherwise. *Post-Publication (P)* is equal to one if the month is after the official publication date and zero otherwise. *Time* is the number of months divided by 100 post-January 1926. *Post-1993* is equal to one if the year is greater than 1993 and zero otherwise. *1-Month Return* and *12-Month Return* are the predictor's return from the last month and the cumulative return over the last 12 months. Standard errors (in parentheses) are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Time	−0.069***		−0.069***			
	(0.011)		(0.026)			
Post-1993		−0.120	0.303***			
		(0.074)	(0.118)			
Post-Sample			−0.190**	−0.179**	−0.132*	−0.128
			(0.081)	(0.080)	(0.076)	(0.078)
Post-Publication			−0.362***	−0.310**	−0.295***	−0.258***
			(0.124)	(0.122)	(0.089)	(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

Do Predictor Types Vary?

Four Predictor Types

1. Event - share issues, changes in analyst recommendations, R&D changes
2. Market - volume, price, returns, shares outstanding, momentum
3. Valuation - sales-to-price, book-to-market
4. Fundamentals - Debt, taxes, accruals

Do Predictor Types Vary?

Variable	(1)	(2)	(3)	(4)
Post-Publication (P)	−0.208***	−0.316***	−0.310***	−0.301***
	(0.059)	(0.097)	(0.080)	(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***	0.606***	0.585***	0.630***
	(0.036)	(0.052)	(0.000)	(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

Costly Arbitrage

Predictor portfolios with stocks that are costlier to arbitrage should decline less post-publication.

Hypothesis

If predictor returns are the outcome of rational asset pricing, then the post-publication decline should not be related to arbitrage costs such as:

1. Size
2. Spreads
3. Dollar Volume
4. Idio. Risk
5. Dividend

Additionally, use first principle component of all 5.

Costly Arbitrage

Method

- For each month compute average cross-sectional ranking for a trait between 0 and 1
- Estimate average rank of the stocks in the long or short sides of the portfolio
- Creates monthly average time series for each trait
- Take average of each time series to estimate single costly arbitrage predictor

$$R_{i,t} = \alpha_{i,} + \beta_1 \, Post\ Publication\ Dummy_{i,t} + \beta_2 \, Arbitrage\ Cost_i + \beta_3 \, Post\ Publication\ Dummy_{i,t} \times Arbitrage\ Cost_i + e_{it},$$

Costly Arbitrage

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Post-Pub. (P)	−0.190	−0.139	0.215	−0.242	−0.321	−0.264**
	(0.274)	(0.235)	(0.230)	(0.273)	(0.211)	(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***	0.146*	0.476***	−0.469***	0.855***	0.565***
	(0.130)	(0.174)	(0.144)	(0.171)	(0.097)	(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-Publication Trading Activity in Predictor Portfolios

Hypothesis

If academic publication provides market participants with information, then informed trading activity should affect not only prices, but other indicators of trading

Table VI

Variables	Variance	Trading volume	Dollar volume	Short-long short interest
Post-Sample (S)	−0.054***	0.092***	0.066***	0.166***
	(0.007)	(0.001)	(0.007)	(0.014)
Post-Publication (P)	−0.065***	0.187***	0.097***	0.315***
	(0.008	(0.013)	(0.007)	(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

Conclusion

Summary

- 97 predictors formed in extreme quintiles test for pre-, post-, sample and post-publication periods
- There is some statistical bias, capped at 26%
- Average predictors return declines by 58% post-publication (32% after bias)
- There appears to be decay, and results due to mispricing
- Costlier arbitrage has higher returns - attracts sophisticated investors

Other Points

- Here we look at averages, but this could highlight a few important factors and diminish others
- No behavioral anomalies here