

HOU, XUE, & ZHANG: Replication Controversies in Finance & Accounting

Posted on [14th June 2017](#) by replicationnetwork

[Leave a Comment](#)

[NOTE: This entry is based on the article “Replicating Anomalies” (SSRN, updated in June 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2961979)

Finance academics have started to take replication studies seriously. As hundreds of factors have been documented in recent decades, the concern over p-hacking has become especially acute. In a pioneering meta-study, Harvey, Liu, and Zhu (2016) introduce a multiple testing framework into empirical asset pricing. The threshold t -cutoff increases over time as more factors have been data-mined. A new factor today should have a t -value exceeding 3.

Reevaluating 296 significant factors in published studies, Harvey et al. report that 80-158 (27%-53%) are false discoveries. Two publication biases are likely responsible for the high percentage of false positives. First, it is difficult to publish a negative result in top academic journals. Second, more subtly, it is difficult to publish replication studies in finance, while in many other disciplines replications routinely appear in top

Search for:



Subscribe via RSS

Tweets by

[@TheReplicateNet](#)



Replication Net...
[@TheReplicateN...](#)

At this exponential rate of growth, there will soon be more journals offering Registered Reports than there are journals!

Nov 4, 2017

Replication Network
Retweeted



Brian Nosek
[@BrianNosek](#)

Five short videos from [@SPSPnews](#) on preregistration & RRs, power, replication, & open data from open science stars!
youtube.com/watch?v=1rFWeT...



Nov 2, 2017



Replication Net...
[@TheReplicateN...](#)

journals. As a result, finance and accounting academics tend to focus on publishing new factors rather than rigorously verifying the validity of published factors.

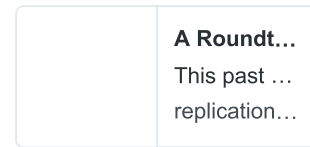
Harvey (2017) elaborates the complex agency problem behind the publication biases. Journal editors compete for citation-based impact factors, and prefer to publish papers with the most significant results. In response to this incentive, authors often file away papers with results that are weak or negative, instead of submitting them for publication. More disconcertingly, authors often engage in, consciously or subconsciously, p-hacking, i.e., selecting sample criteria and test procedures until insignificant results become significant. The outcome is an embarrassingly large number of false positives that cannot be replicated in the future.

We conduct a massive replication of the published factors by compiling a largest-to-date data library with 447 variables. The list includes 57, 68, 38, 79, 103, and 102 variables from the momentum, value-versus-growth, investment, profitability, intangibles, and trading frictions categories, respectively. We use a consistent set of replication procedures throughout. To control for microcaps (stocks that are smaller than the 20th percentile of market equity for New York Stock Exchange, or NYSE, stocks), we form testing deciles with NYSE breakpoints and value-weighted returns. We treat a variable as a replication success if its average return spread is significant at the 5% level.

Our replication indicates rampant p-hacking in the published literature. Out of 447 factors, 286 (64%) are insignificant at the 5% level. Imposing the t -cutoff of 3 per Harvey, Liu, and Zhu (2016) raises the number of insignificance to 380 (85%).

The biggest casualty is the liquidity literature. In the trading frictions category, 95 out of 102 variables (93%) are insignificant. Prominent variables that do not survive our replication include Jegadeesh's (1990) short-term reversal;

A roundtable podcast on the merits of lowering the threshold for statistical significance to 0.005
[@JustinEsarey](#)
[replicationnetwork.com/2017/10/30/a-r...](#)



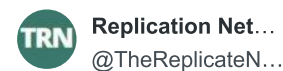
Oct 29, 2017

Replication Network Retweeted



that awkward moment when you and all your friends show up to the same halloween party dressed as $p = .049$

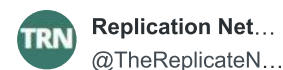
Oct 26, 2017



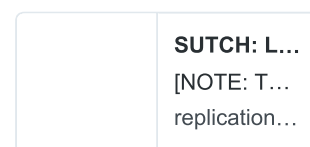
Another Elsevier journal goes "blind" [@dDeWeerd](#)
[replicationnetwork.com/2017/10/27/ano...](#)



Oct 26, 2017



GUEST BLOG at TRN: Lesson learned from replicating Piketty — How not to do economic history [@ATabarrok](#)
[replicationnetwork.com/2017/10/26/sut...](#)



Oct 26, 2017



Datar-Naik-Radcliffe's (1998) share turnover; Chordia-Subrahmanyam-Anshuman's (2001) coefficient of variation for dollar trading volume; Amihud's (2002) absolute return-to-volume; Acharya-Pedersen's (2005) liquidity betas; Ang-Hodrick-Xing-Zhang's (2006) idiosyncratic volatility, total volatility, and systematic volatility; Liu's (2006) number of zero daily trading volume; and Corwin-Schultz's (2012) high-low bid-ask spread. Several recent friction variables that have received much attention are also insignificant, including Bali-Cakici-Whitelaw's (2011) maximum daily return; Adrian-Etula-Muir's (2014) intermediary leverage beta; and Kelly-Jiang's (2014) tail risk.

The much researched distress anomaly is virtually nonexistent. Campbell-Hilscher-Szilagyi's (2008) failure probability, the O-score and Z-score in Dichev (1998), and Avramov-Chordia-Jostova-Philipov's (2009) credit rating all produce insignificant average return spreads.

Other influential but insignificant variables include Bhandari's (1988) debt-to-market; Lakonishok-Shleifer-Vishny's (1994) five-year sales growth; several of Abarbanell-Bushee's (1998) fundamental signals; Diether-Malloy-Scherbina's (2002) dispersion in analysts' forecast; Gompers-Ishii-Metrick's (2003) corporate governance index; Francis-LaFond-Olsson-Schipper's (2004) earnings attributes, including persistence, smoothness, value relevance, and conservatism; Francis et al.'s (2005) accruals quality; Richardson-Sloan-Soliman-Tuna's (2005) total accruals; and Fama-French's (2015) operating profitability, which is a key variable in their 5-factor model.

Even for significant anomalies, their magnitudes are often much lower than originally reported. Famous examples include Jegadeesh-Titman's (1993) price momentum; Lakonishok-Shleifer-Vishny's (1994) cash flow-to-price; Sloan's (1996) operating accruals; Chan-Jegadeesh-Lakonishok's (1996) standardized unexpected

Check out what's recent at
The Replication Network
(October 2017) -
mailchi.mp/a755960fd536/u...

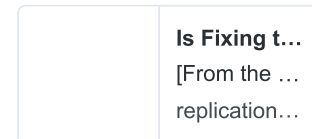
Oct 22, 2017



Replication Net...
@TheReplicateN...

Is fixing the replication crisis as simple as lowering the p-value?

replicationnetwork.com/2017/10/21/is-...



Oct 20, 2017



Replication Net...
@TheReplicateN...

Replicating Piketty's "Capital in the Twenty-First Century"

@ATabarrok

Tags

Andrew Gelman BITSS Brian Nosek Journal Data Policies Journals Nature journal negative results p-hacking p-value

Psychology

publication bias

replication

replication crisis replications

Reproducibility

Reproducibility crisis retraction

Retraction Watch significance

testing Transparency

Blogroll

Berkeley Initiative for

Transparency in the Social Sciences

Curate Science

Political Science Replications

earnings, abnormal returns around earnings announcements, and revisions in analysts' earnings forecasts; Cohen-Frazzini's (2008) customer momentum; and Cooper-Gulen-Schill's (2008) asset growth.

Why does our replication differ so much from the original studies? The key word is microcaps. Microcaps represent only 3% of the total market capitalization of the NYSE-Amex-NASDAQ universe, but account for 60% of the number of stocks. Microcaps not only have the highest equal-weighted returns, but also the largest cross-sectional standard deviations in returns and anomaly variables. Many studies overweight microcaps with equal-weighted returns, and often together with NYSE-Amex-NASDAQ breakpoints, in portfolio sorts.

Hundreds of studies use cross-sectional regressions of returns on anomaly variables, assigning even higher weights to microcaps. The reason is that regressions impose a linear functional form, making them more susceptible to outliers, which most likely are microcaps. Alas, due to high costs in trading these stocks, anomalies in microcaps are more apparent than real. More important, with only 3% of the total market equity, the economic importance of microcaps is small, if not trivial.

Our low replication rate of only 36% is not due to our extended sample relative to the original studies. Repeating our replication in the original samples, we find that 293 (66%) factors are insignificant at the 5% level, including 24, 44, 13, 38, 81, and 93 across the momentum, value-versus-growth, investment, profitability, intangibles, and trading frictions categories, respectively. Imposing the t -cutoff of three raises the number of insignificance to 387 (86.6%). The total number of insignificance at the 5% level, 293, is even higher than 286 in the extended sample. In all, the results from the original samples are close to those from the full sample.

[Replication Wiki](#)

[Retraction Watch](#)

[Statistical Modeling, Causal Inference, and Social Science](#)

Recent Posts

[MCSHANE & GAL: Statistical Significance and Dichotomous Thinking Among Economists](#)

[A Roundtable Podcast on the Merits of Lowering the Threshold for Statistical Significance to 0.005](#)

[Another Elsevier Journal Goes "Blind"](#)

[SUTCH: Lesson Learned from Replicating Piketty — How Not To Do Economic History](#)

[Is Fixing the Replication Crisis As Simple as Lowering the p-Value?](#)

We also use the Hou, Xue, and Zhang (2015) q -factor model to explain the 161 significant anomalies in the full sample. Out of the 161, the q -factor model leaves 115 alphas insignificant (150 with $t < 3$). In all, capital markets are more efficient than previously recognized.

Kewei Hou is Fisher College of Business Distinguished Professor of Finance at The Ohio State University. Chen Xue is Assistant Professor of Finance at University of Cincinnati. Lu Zhang is the John W. Galbreath Chair, Professor of Finance, at The Ohio State University. Correspondence about this blog should be addressed to Lu Zhang at zhanglu@fisher.osu.edu.

REFERENCES

Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal returns to a fundamental analysis strategy. *The Accounting Review*, 73, 19-45.

Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics*, 77, 375-410.

Adrian, T., Etula, E., & Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, 69, 2557-2596.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time series evidence. *Journal of Financial Markets*, 5, 31-56.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61, 259-299.

Avramov, D., Chordia, T., Jostova, G., & Philipov, A. (2009). Credit ratings and the cross-section of stock returns. *Journal of Financial Markets*, 12, 469-499.

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Mxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99, 427-446.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *Journal of Finance*, 43, 507-528.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *Journal of Finance*, 63, 2899-2939.

Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies, *Journal of Finance*, 51, 1681-1713.

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, 59, 3-32.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns, *Journal of Finance*, 63, 1977-2011.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset growth and the cross-section of stock returns, *Journal of Finance*, 63, 1609-1652.

Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance*, 67, 719-759.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, 1, 203-219.

Dichev, I. (1998). Is the risk of bankruptcy a systematic risk? *Journal of Finance*, 53, 1141-1148.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns, *Journal of Finance*, 57, 2113-2141.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model, *Journal of Financial Economics*, 116, 1-22.

Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2004). Cost of equity and earnings attributes, *The Accounting Review*, 79, 967-1010.

Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2005). The market price of accruals quality, *Journal of Accounting and Economics*, 39, 295-327.

Gompers, P., Ishii, J., & Metrick, A. (2001). Corporate governance and equity prices, *Quarterly Journal of Economics*, 118, 107-155.

Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *Journal of Finance*, forthcoming.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ...and the cross-section of expected returns. *Review of Financial Studies*, 29, 5-68.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28, 650-705.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45, 881-898.

Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48, 65-91.

Kelly, B., & Jiang, H. (2014). Tail risk and asset prices. *Review of Financial Studies*, 27, 2841-2871.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk, *Journal of Finance*, 49, 1541-1578.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, 82, 631-671.

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices, *Journal of Accounting and Economics*, 39, 437-485.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71, 289-315.

Share this:



Be the first to like this.

☰ Category: GUEST BLOGS 🔖 Tags: Accounting, Anomalies, Finance, Lu Zhang, replication

← [Reasons for Loving Null Results](#)

[Calling All PhD Students: Do a Replication. Get Paid.](#) →

Leave a Reply

Enter your comment here...