Does Academic Research Destroy Stock Return Predictability?*

R. David McLean
University of Alberta
Phone: 774-270-2300
Email: rdmclean@ualberta.ca

Jeffrey Pontiff
Boston College
Phone: 617-552-6786
Email: pontiff@bc.edu

March 24, 2014

Abstract
We study the out-of-sample and post-publication return-predictability of 95 characteristics that are shown in published academic studies to predict cross-sectional stock returns. We estimate an out-of-sample decline in predictability due to statistical bias of 13%, which is not statistically significant. The average post-publication decline, which we attribute to both statistical bias and informed trading, is about 44%, and statistically different from both 0% and 100%. Our findings point to mispricing as the source of predictability. Predictors with larger in-sample returns and predictors that potentially violate weak-form market efficiency decline more post-publication. Predictor portfolios that are easier to arbitrage, i.e., consisting more of stocks with low idiosyncratic risk and high liquidity, also decline more post-publication. Post-publication, predictor portfolios have increases in short interest on the short side, and increases in correlations with other portfolios that are based on published predictors.

Keywords: Return predictability, limits of arbitrage, publication impact, market efficiency, comovement, statistical bias.

JEL Code: G00, G14, L3, C1

---

* We are grateful to the Q Group and the Dauphine-Amundi Chair in Asset Management for financial support. We thank participants at the Financial Research Association's 2011 early ideas session, seminar participants at Babson College, Bocconi University, Brandeis University, Boston College, CKGSB, HBS, Georgia State University, HEC Montreal, MIT, Northeastern University, University of Toronto, University of Maryland, City University of Hong Kong International Conference, Finance Down Under Conference 2012, University of Georgia, University of Washington Summer Conference, European Finance Association (Copenhagen), 1st Luxembourg Asset Management Conference, and Pierluigi Balduzzi, Turan Bali, David Chapman, Mark Bradshaw, Shane Corwin, Alex Edmans, Lian Fen, Wayne Ferson, Francesco Franzoni, Xiaohui Gao, Thomas Gilbert, Robin Greenwood, Bruce Grundy, Clifford Holderness, Darren Kisgen, Owen Lamont, Jay Ritter, Andrei Shliefer, Paul Schultz, Bruno Skolnik, Jeremy Stein, Matti Suominen, Allan Timmermann, Michela Verado, Artie Woodgate, and Jianfeng Yu for helpful conversations.

Finance research has uncovered many cross-sectional relations between predetermined variables and future stock returns. Beyond historical curiosity, these relations are relevant to the extent they provide insight into the future. Whether or not the typical relation continues outside of a study's original sample is an open question, the answer to which can shed light on why cross-sectional return predictability is observed in the first place.[1] Although several papers note whether a specific cross-sectional relation continues, no study compares in-sample returns, post-sample returns, and post-publication returns among a large sample of predictors. Moreover, previous studies produce contradictory messages. As examples, Jegadeesh and Titman (2001) show that the relative returns to high-momentum stocks increased after the publication of their 1993 paper, while Schwert (2003) argues that since the publication of the value and size effects, index funds based on these variables fail to generate alpha.[2]

In this paper, we synthesize information from 95 predictors that have been shown to explain cross-sectional stock returns in peer-reviewed finance, accounting, and economics journals. Our goal is to better understand what happens to return-predictability outside of a study's sample period. We compare each predictor's returns over three distinct periods: (i) the original study's sample; (ii) after the original sample but before publication; and (iii) post-publication. Previous studies contend that return-predictability is either the outcome of a rational asset pricing model, statistical biases, or mispricing. By comparing return-predictability across

[1] We focus on cross-sectional variables. For an analysis of the performance of time-series variables, see LeBaron (2000) and Goyal and Welch (2008). For an analysis of calendar effects, see Sullivan, Timmermann, and White (2011).

[2] Lewellen (2011) uses 15 variables to produce a singular rolling cross-sectional return proxy and shows that it predicts, with decay, next period's cross section of returns. Haugen and Baker (1996) and Chordia, Subrahmanyan, and Tong (2013) compare characteristics in two separate subperiods. Haugen and Baker show that each of their characteristics produces statistically significant returns in the second-subperiod, whereas Chordia, Subrahmanyan, and Tong show that none of their characteristics is statistically significant in their second-subperiod. Green, Hand, Zhang (2012) identify 300 published and unpublished characteristics but they do not estimate characteristic decay parameters as a function of publication or sample-end dates.

these three distinct periods, we are able to give insight into what best explains the typical predictor's returns.

Pre-publication, out-of-sample predictability. If return-predictability in published studies is the result of statistical biases, then predictability should disappear out of sample. We use the term "statistical biases" to describe a broad array of biases that are inherent to research.

At least three statistical biases could affect observed stock return-predictability: sample selection biases, method selection biases, and multiple testing bias. Leamer (1978) shows the impact of "specification search" biases, which occur if the choice of method is influenced by the methodology's result. Lo and MacKinlay (1990) examine a specific type of specification search bias found in finance, which they refer to as the "data snooping bias." A second type of bias is sample selection bias, studied in Heckman (1979), where the sample construction is influenced by the result of the test. [3] A third type of bias arises when researchers conduct multiple tests of the same hypothesis. This bias goes back to Bonferroni (1935) and is applied to finance by Fama (1991) when he notes that, "With clever researchers on both sides of the efficiency fence, rummaging for forecasting variables, we are sure to find instances of 'reliable' return predictability that are in fact spurious." Harvey, Liu, and Zhu (2013) argue that this bias has worsened over time; as researchers test an increasing number of predictor's, an increasing number of studies will be published with spurious predictability. To the extent that the results of the studies in our sample are caused by such biases, we should observe a decline in return-predictability out-of-sample.

Post-publication predictability. We assume that more market participants know about a predictor after a paper documenting the predictor has been published, as compared to before the

[3] Along these lines, a strategy's spuriously high returns can attract academic attention to the strategy, making the publication date endogenous. We thank Allan Timmermann for pointing out this possibility.

paper's publication date. However, we do not assume that the publication date is a precise transition date. Papers are often presented at conferences and distributed before publication, causing information to be released before the publication date. On the other hand, market participants may be slow to respond to academic studies, so information may begin to work its way into prices long after the publication date. Most of our tests therefore examine whether return-predictability is different after the publication date as compared to before the publication date. [4]

Whether return predictability should change post-publication or not depends on what causes the predictability. Cochrane (1999) explains that if predictability reflects risk, then it is likely to persist regardless of how many people know about it: "Even if the opportunity is widely publicized, investors will not change their portfolio decisions, and the relatively high average return will remain." Cochrane's logic follows Muth's (1961) rational expectations hypothesis, and thus can be broadened to non-risk models such as Amihud and Mendelson's (1986) transaction-based model and Brennan's (1970) tax-based model. If return-predictability reflects rational expectations, then pre- and post-publication (and out-of-sample but pre-publication) return-predictability should be similar.

Alternatively, if return-predictability is the result of mispricing and if publication draws the attention of sophisticated investors who trade against the mispricing, then we would expect the effects to disappear after the paper is published or at least continue at a reduced level if costs prevent arbitrage from fully eliminating mispricing (see Delong, Shleifer, Summers, and

[4] To the best of our knowledge, the first empirical examination of the effects of academic research on capital markets is Mittoo and Thompson's (1990) study of the size effect. They use a regime switching model to illustrate a post-1983 difference in returns to size portfolios.

Waldman (1990), Pontiff (1996, 2006), and Shleifer and Vishny (1997)).[5] We can differentiate this effect from that of statistical biases by finding a greater decline post-publication as compared to any decline out of-sample but pre-publication.

Findings. We conduct our analysis using 95 different characteristics from 68 different studies. Using long-short portfolio strategies that buy and sell extreme quintiles that are based on each predictor, we confirm significant returns for 78 of the characteristics in-sample; for 16 of the characteristics we could not, using our portfolio method, find statistically significant return predictability in the original sample. We focus our analyses on the 78 characteristics with significant in-sample returns. As we mention above, the post-sample but pre-publication period is useful for estimating statistical bias. We find that on average, return-predictability declines by 13% during this period, implying that a 1% in sample return reflects a bias-free in sample return of 0.87%. This finding is statistically insignificant—we cannot reject the hypothesis that there are no statistical biases. Our 13% estimate is probably too high, since some traders could learn about the predictor before publication and their actions will cause some decay that is captured in the 13%.

We estimate that the average predictor's return decays 44% post-publication. Thus, an in-sample return of 1% is expected to decay to 0.56% post-publication. Combining this finding with an estimated statistical bias of 13% implies a lower bound on the publication effect of about 31%. We can reject the hypothesis that post-publication return-predictability does not change and we can also reject the hypothesis that return-predictability disappears entirely. The post-publication decline is robust to time indicators used by other authors and time fixed effects.

[5] For evidence of limited arbitrage in short sellers and mutual funds, see Duan, Hu, and McLean (2009 and 2010).

The decay in predictability is larger for predictors with higher in sample returns and higher in sample t-statistics. We also find evidence that decay is larger for predictors that can be constructed with only price and trading data, and therefore represent violations of weak form market efficiency. The post-publication decline is greater for predictors that are less costly to arbitrage; i.e., predictors that require more trading in stocks with high liquidity and low idiosyncratic risk. Our findings are consistent with mispricing explanations for return-predictability, as post-publication returns decline the most for portfolios that have the highest in-sample returns and lowest arbitrage costs.

We further investigate the effects of publication by studying traits that reflect trading activity. We find that stocks within the predictor portfolios have post-publication increases in variance, turnover, and dollar volume. The difference in the relative amount of short interest between stocks in the short and long sides of each portfolio also increases after publication. These findings are consistent with the idea that academic research draws attention to predictors.[7]

The correlation across predictors is quite low, averaging only 4.9%. We find that correlations between predictors are affected by publication. We find that yet-to-be-published predictor portfolios are correlated. However, after a predictor is published its correlation with other yet-to-be-published predictor portfolios decreases, while its correlation with other already-published predictor portfolios increases. One interpretation of this finding is that predictors are the result of mispricing and mispricing has a common source; this is why in-sample predictor portfolios are correlated. This interpretation is consistent with the irrational comovement models proposed in Lee, Shleifer, and Thaler (1991) and Barberis and Shleifer (2003). Publication then causes more arbitrageurs to trade on the predictor, which causes predictor portfolios to become

[7] Drake, Rees and Swanson (2011) demonstrate that short interest is more pronounced in the low-return segment of several characteristic sorted portfolios. Their study does not account for the difference between in- and out-of-sample short interest.

more correlated with already-published predictor portfolios that are also pursued by arbitrageurs, and less correlated with yet-to-be-published predictor portfolios.

1. Research Method

We identify studies that find cross-sectional relations between variables that are known in a given month and stock returns in the following month(s). We do not study time series predictability. We limit ourselves to studies in academic peer-reviewed finance, accounting, and economics journals, where the null of no cross-sectional predictability is rejected at the 5% level, and to studies that can be constructed with publicly available data. Most often, these studies are identified with search engines such as Econlit by searching for articles in finance and accounting journals with words such as "cross-section." Some studies are located from reference lists in books or other papers. Lastly, in the process of writing this paper, we contacted other finance professors and inquired about cross-sectional relations that we may have missed.

Most studies that we identify either demonstrate cross-sectional predictability with Fama-MacBeth (1973) slope coefficients or with long-short portfolio returns. Some of the studies that we identify demonstrate a univariate relation between the characteristic and subsequent returns, while other studies include additional control variables. Some studies that we identify are not truly cross-sectional, but instead present event-study evidence that seems to imply a cross-sectional relation. Since we expect the results from these studies to provide useful information to investors, we also include them in our analyses.

We use 95 cross-sectional relations from 78 different studies. We include all variables that relate to cross-sectional returns, including those with strong theoretical motivation such as Fama and MacBeth's landmark 1973 study of market beta in the *Journal of Political Economy* and

Amihud's 2002 study of a liquidity measure in the *Journal of Financial Markets*. The study with the most number of original cross-sectional relations that we utilize (4) is Haugen and Baker's 1996 study in the *Journal of Financial Economics*. Haugen and Baker (1996) investigate more than four predictors, but some of their predictors were documented by other authors earlier and are therefore associated with other publications in our study.

We are unable to exactly construct all of the characteristics. In such cases, we calculate a characteristic that captures the intent of the study. As examples, Franzoni and Marin (2006) show that a pension funding variable predicts future stock returns. This variable is no longer covered by Compustat, so we use available data from Compustat to construct a variable that we expect to contain much of the same information. Dichev and Piotroski (2001) show that firms that are downgraded by Moody's experience negative future abnormal returns. Compustat does not cover Moody's ratings. It does cover S&P ratings, so we use S&P rating downgrades instead. Characteristics that use accounting data are winsorized, such that values that are below the 1^{st} percentile are assigned the value of the 1^{st} percentile, and values that are above the 99th percentile are assigned the value of the 99^{th} percentile. Returns are equally weighted unless the primary study presents value-weighted portfolio results (e.g., Ang, Hodrick, Xing, and Zhang, 2006).

We estimate each predictors' return-predictability by computing the return of a portfolio that each month invests in stocks in the top 20^{th} percentile of the characteristic (the strategy's long-side) minus the return of a portfolio that invests in stocks in the bottom 20^{th} percentile of the characteristic. In an earlier version of the paper we also calculated monthly Fama-MacBeth (1973) slope coefficient estimates using a continuous measure of the characteristic (e.g. firm size or past returns). As Fama (1976) shows, Fama-MacBeth slope coefficients are returns from long-short portfolios with unit net exposure to the characteristic. We obtained similar findings using both methods, so for the sake of brevity we only report quintile returns.

We segment periods based on the end of the sample and the publication date because they are clear, agreeable dates that may be associated with changes in predictability. The end of the original sample provides a clear point to estimation statistical bias. The publication date however, is not a clear point for examining the impact of market participants learning about a predictor (assuming the predictor is the result of mispricing). As we mention above, we assume that more investors know about a predictor during the sample period after the publication date as compared to the sample period before the publication date. Some market participants will read a working paper version before publication, while some will read the paper years after publication. Hence, post-publication decay in return-predictability may be a slow process and we are unaware of theories of how long the decay should take and the functional form of the decay. Despite the simplicity of our approach, the publication date generates robust estimates of return decay.

2. Creating the Data and In-Sample Replicability

Summary statistics for the characteristics that we study are provided in Table 1. Our goal is not to perfectly replicate a paper. This is impossible since CRSP data changes over time and papers often omit details about precise calculations. Seventeen of our predictors produce in-sample t-statistics that are between -1.50 and 1.50.[8] We do not include these characteristics in the paper's main tests. Thus, a total of 78 (95 – 17) or 82% of the predictor's produce significant in-sample returns and are used in the paper's primary tests.

[8] If a characteristic is not associated with a t-statistic outside of the -1.50 to 1.50 range, both co-authors independently wrote code to estimate the effect.

As we mention above, in an earlier version of the paper we also estimated predictor returns using continuous variables in Fama-MacBeth regressions. We are able to find significant in-sample returns for three additional predictors using this method. One might therefore claim that of the 95 predictors, we are able to replicate in-sample predictability for 81 or 85% of them.

Admittedly, the decision to use a t-statistic cut-off of 1.50 is arbitrary. The decision was motivated by a desire to utilize as many characteristics as possible, while still measuring the same essential characteristic as the original paper. Given that some papers feature characteristics with t-statistics that are close to 2.0 and that we are not perfectly replicating the original authors' methodology, a cut-off of 1.50 seemed reasonable to us. That stated, only two of the 78 characteristics that we include in the paper's analyses only four have t-statistics that are less than 1.80.

We define the publication date as the date based on the journal's year and issue. For this date convention, the average length of time between the end of the sample and publication is 55 months. For comparison, the average original in-sample span is 329 months, and the average out-of-sample span is 141 months. In an earlier version of the paper we also consider the publication date to be the earlier of the actual publication date and the first time that paper appeared on the SSRN. The average number of months between the end of the sample and SSRN date is 44 months, and we get the same findings using this method. Table 1 shows that the generous number of cross-sectional characteristics yields a total number of 4,361 out-of-sample but pre-publication observations, and 11,145 post-publication observations.

3. Main Results

3.1. Characteristic Dynamics Relative to End of Sample and Publication Dates

We now formally study the returns of each predictor relative to its sample-end and publication dates. The baseline regression model is described in equation (1):

$$R_{it} = \alpha_{i,} + \beta_1 \, Post\ Sample\ Dummy_{i,t} \ + \ \beta_2 \, Post\ Publication\ Dummy_{i,t} \ + \ e_{it} \qquad (1)$$

In equation 1 the dependent variable is the monthly return for predictor *i* in month *t*. the post-sample dummy is equal to one if month *t* is after the end of the original sample but still pre-publication and zero otherwise, while the post-publication dummy is equal to 1 if the month is post-publication and zero otherwise. The variable α_i is the predictor's fixed effect. We report the average value of α_i as the intercept in the tables.

As we mention previously, correlations across predictors are low, averaging only 4.9%. Nonetheless, to be conservative we estimate equation (1) using generalized least squares, which accounts for correlations across predictors when estimating the coefficients. That stated, using simpler ordinary least squares regressions produces the same findings. In addition, we also also cluster our standard errors on time to control for any common shocks across predictors. In unreported results we cluster on anomaly, and produce larger t-statistics.

The post-sample coefficient estimates the total impact of statistical biases on predictor in sample performance (under the assumption that sophisticated traders are unaware of the working paper before publication). The post-publication coefficient estimates both the impact of statistical biases and the impact of publication. If statistical biases are the cause of in-sample predictability, then the coefficients for both the post-sample and the post-publication dummy should be equal to -1. Such a finding would be consistent with Fama's (1991) conjecture that return-predictability in academic studies is the outcome of data-mining.

Instead, if predictors' returns are the result of mispricing and arbitrage resulting from publication corrects all mispricing, then the post-publication coefficient will be equal to -1 and the post-sample dummy will be close to zero. In the other extreme, if there are no statistical biases and academic papers have no influence on investors' actions, then both of the coefficients should equal zero.

3.2. Predictor Return Dynamics Relative to End of Sample and Publication Dates

Table 2 presents regression estimates of how predictability varies through the life cycle of a publication. Column 1 reports the results for our baseline specification, which is an estimate of Equation 1 within the sample of the 79 predictors for which we found significant in sample predictability. The post-sample coefficient in this regression is -0.104, and statistically insignificant. This means that our best estimate of the post-sample decline 10.4 basis points, although the standard errors are too large to claim significance at standard levels. The post-publication coefficient is -0.351, and it is significant at the 1% level. This means that on average predictor portfolios are 35.1 basis points lower post-publication compared to before publication. The intercept in this regression is 0.795, and it is also significant at the 1% level. The reported intercept here is the average of the predictor dummies. The intercept therefore shows that the average predictor has an in-sample mean return of 79.5 basis points per month. Hence, the post-sample and post-publication declines relative to the in-sample mean are 44% and 13% respectively.

The regression in the second column includes all 95 predictors, and therefore includes the 16 predictors for which we did not find significant in-sample predictability. The inclusion of these additional predictors does not change the basic inference reported in column 1. The post-sample and post-publication coefficients are -0.091 and -0.300 respectively in column 2, similar to the results in column 1 and with the same levels of statistical significance as in column 1. The intercept in column 2 is 0.683, similar to but slightly smaller than the intercept reported in column 1. The intercept and coefficients are expected to be smaller in column 2, as we are including 16 predictors that do not have significant in-sample predictability.

In the regression reported in third column we exclude the predictor dummies and in their place include the in-sample mean of each predictor as an independent variable. With respect to the post-sample and post-publication coefficients, they are both similar to the coefficients in the first column, and show that there is an insignificant decline of about 10 basis points out-of-sample, and then a significant decline of 32 basis points post-publication. The coefficient for the in-sample mean is 0.912, and it is highly significant. This means that average predictor's returns is about 9% lower post-publication as compared to its in-sample mean. This also shows that predictors with higher in-sample means have larger declines post-publication in absolute terms.

At the bottom of Table 3, we report tests of whether the coefficient for post-publication is greater than the coefficient for out-of-sample but pre-publication. In all three regressions that we describe above the difference is significant. Hence, the decline in return-predictability that is observed post-publication is significantly higher than the decline in return-predictability that is observed out-of-sample, but pre-publication. This difference tells us that there is an effect associated with publication that cannot be explained by statistical biases, which should be fully reflected in the out-of-sample but pre-publication coefficients.

The fourth column returns to the regressions that include predictor dummy variables. In these regressions we include interactions between the in-sample mean return of each predictor and the out-of-sample and post-publication dummy variables. The interactions test whether predictors with higher in-sample means decline more post-publication. We do not include the in-sample mean in the regression by itself because it does not vary over time and we include predictor dummies. If predictor dummies are included, then only variables that vary within the predictor over time can have effects on the regression's dependent variable.

In column 4 the coefficient for post-sample is 0.088 and insignificant, while the coefficient for the post-sample interaction with the in-sample mean is -0.247. The mean predictor has an in-sample mean return of 0.785 (see Table 1), so the overall post-sample effect is 0.088 + (-0.247 x 0.785) = -0.106, similar to the post-sample coefficient in column 1. The in-sample mean has a standard deviation of 0.562. (see Table 1) Hence, a one standard deviation increase in the in-sample mean leads to an additional -0.247 x 0.562 =13.9 basis point decline in predictor returns post-sample. This could reflect the fact that predictors with larger in-sample returns are likely to have a higher degree of statistical bias. Alternatively, it could reflect the fact that arbitrageurs or more likely to learn about and trade on predictors with higher returns before publication.

The post-publication coefficient in column 4 is -0.070 and insignificant, while the post-publication interaction is -0.342 and highly significant. The average predictor therefore has a post-publication decline of -0.070 + (-0.342 x 0.785) = -0.268, which is similar to the effect estimated in column 1. A one standard deviation increase in the in-sample mean leads to an additional -0.342 x 0.562 =19.4 basis point decline in post-publication returns. This relation is also displayed in Figure 1.A, which plots the average in-sample mean for each predictor against its post-publication decline, and shows that predictors with larger in-sample returns have greater post-publication declines. This finding is consistent with the idea that arbitrageurs are more likely to trade on predictors that offer higher returns, and that publication advertises predictors to investors.

The final regression in Table 2 interacts the post-sample and post-publication dummies with the predictor's in-sample t-statistic. The post-sample coefficient in this regression 0.027, while the post-sample-t-statistic interaction coefficient is -0.029. Both coefficients are insignificant. The in-sample t-statistics have a mean of 4.46 and a standard deviation of 2.96 (not reported in tables). Hence, the regression estimates an average decline of 10.2 basis points post-sample, similar to what is reported in in column 1. A one standard deviation increase in the in-sample t-statistic leads to an additional decline of 8.60 basis points.

The post-publication coefficient in column 5 is -0.160 and insignificant, while the post-publication interaction coefficient is -0.044. Hence, for the mean predictor its post-publication decline is about 36 basis points. A one standard deviation increase in the in-sample t-statistic leads to an additional 13 basis points in decline. This relation is plotted in Figure 1.B, which displays the relation between the in-sample t-statistic and the post-publication decline in returns, and shows that predictors with larger in-sample t-statistics have larger post-publication declines. The results here are consistent with the idea that arbitrageurs target predictors with more reliable returns.

3.3. A Closer Look at Predictor Return Dynamics around the Sample-End and Publication Dates

Figure 2 further considers changes in predictability by examining finer post-sample and post-publication partitions. The figure plots the coefficients from a regression of predictor returns on dummy variables that signify the last 12 months of the original sample; the first 12 months out-of sample; and the other out-of-sample months. In addition, the publication dummy is split up into six different variables; one dummy variable for each of the first five years post-publication, and a dummy variable for all of the months that are at least five years after publication.

The publication process often takes years. This gives researchers the opportunity to choose where to end their samples with the purpose of reporting stronger results. Figure 2 shows that the coefficient for the last 12 months of the sample period is positive. This shows that the last 12 months of the sample has higher returns than the other in-sample months, which could be consistent with choosing to end samples opportunistically. However, the coefficient for the first 12 months post-sample is also positive, showing that the first 12 months post-sample has on average higher returns than the average returns in-sample; if authors where selectively choosing their sample periods, then this coefficient should be negative.

Figure 2 shows that after the first 12 months out-of-sample, returns are lower as compared to in-sample, and stay that way throughout the life of the predictor. After the first year post-sample and during the remaining months out-of-sample but before publication, returns are lower by 18 basis points. Returns remain at this level throughout the first two years post-publication, and then begin to decay further. In the third year we estimate a decay of 26 basis points; in the fourth year it is 50.5 basis points; and in the fifth year it is 27 basis points. After the fifth year predictors' returns are on average 39 basis points lower as compared to in-sample.

Taken together what the results in Figure 3 show is that this study is not an event study; nothing magical happens on the day a paper is published. What is shown is that the decline in returns is not in full effect until a few years after the publication data. This suggests it takes time for investors to learn about and effectively implement the strategies described in academic studies.

3.4. Do Predictor Returns Exhibit Time Trends and Persistence?

It could be the case that the dissemination of academic research has no effect on return-predictability, and that our end-of-sample and publication coefficients reflect a time trend or a trend that proxies for lower costs of corrective trading. For example, anomalies might reflect mispricing and declining trading costs have made arbitrage less costly, which is why we observe the drop post-publication. Goldstein, Irvine, Kandel, and Wiener (2009) present evidence that brokerage commissions dropped dramatically from 1977 to 2004, while Anand, Irvine, Puckett and Venkataraman (2012) show that, over the last decade, execution costs have fallen. Chordia, Subrahmanyam, and Tong (2013) show that the returns of the different predictors decline after 1993, and effect they attribute to more hedge funds and lower trading costs. Hence, it could be the case that characteristics are diminishing because the costs of trading on these characteristics have declined over time.

We study these effects in Table 3. We construct a time variable that is equal to 1/100 in January 1926 and increases by 1/100 during each consecutive month in our sample. In column 1 we estimate a regression of monthly portfolio returns on the time variable and a predictor fixed effect. The time variable produces a negative slope coefficient that is significant at the 1% level, which is consistent with the idea that portfolio returns have declined over time. In Column 2, we replace the time variable with three different time variables, each constructed specifically for each predictor. *I-Time* is a time variable that begins during the first month of the in-sample period, and is equal to zero after predictor's in-sample period ends. *S-Time* begins when the predictor 's sample period ends, and ends on the predictor's publication date. Finally, *P-Time* begins on the predictor's publication date and goes through the end of our sample period (2012).

The regression reported in column 2, which use the three predictor-specific time variables, shows that there is not a general time decline in returns throughout the life of a predictor. Instead, there is only a downward trend in returns *after* the predictor has been featured in a published study. The coefficients for the in-sample and out-of-sample trends are both insignificant, whereas there is a negative and significant trend in returns during the post-publication period.

In column 3 we use a dummy variable that is equal to 1 if the year is after 1993 and zero otherwise. We use this specification because, as we mention above, Chordia, Subrahmanyam, and Tong (2013) show that 12 predictors have lower returns post-1993, which is the second half of their study's sample period. The post-1993 coefficient is insignificant in our sample.

Given the results in the first three columns, we estimate a specification in column 4 that includes the dummy variables for post-publication and post-sample along with the three time variables used in column 2. The regressors are correlated so there is a good deal of multicollinearity, which inflates the standard errors. The regression still estimates a sizeable and significant post-publication decline of 25 basis points. The coefficient for the post-publication trend is negative but insignificant. The post-sample dummy is positive but insignificant, while the post-sample trend is negative and significant, suggesting that predictor returns are lower towards the end of the post-sample period, consistent with the results displayed in Figure 3.

Perhaps a cleaner way to control for time effects while studying the effect of publication is to include time fixed effects, as we avoid the collinearity between the post-sample and post-publication dummies and the time variables. We estimate this specification in column 5. This regression estimate an insignificant decline of 11.6 basis points out-of-sample, and a significant 22.4 basis points post-publication. These results therefore show that the post-publication decay is robust to controlling for decay over time.

In the final two regressions in Table 3 we test whether predictor returns are persistent, and whether controlling for persistence changes the publication effect. Recent work by Moskowitz, Ooi, and Pedersen (2010) and Asness, Moskowitz and Pedersen (2009) finds broad momentum across asset classes and correlation of momentum returns across classes. The pervasiveness of the results in these papers suggests that momentum, or perhaps shorter-term persistence, might exist among our larger sample of characteristics.

We include the predictor's last month's return and the sum of its last 12 months' returns in regressions 6 and 7 respectively. Both of the lagged return coefficients are positive and significant, which is broadly consistent with the findings of Moskowitz, et al. The publication coefficient remains significant in each of these regressions, suggesting a post-publication decline of about 30 basis points once past returns are considered.

3.5. Does the Post-Publication Decline Vary Across Predictor Types?

In this section of the paper we ask whether in-sample predictability and post-publication declines vary across predictors, based on the information that is needed to construct the predictor. To conduct this exercise we split our predictors into four groups: (i) Event; (ii) Market; (iii) Valuation; and (iv) Fundamentals.

Event predictors are based on events within the firm, external events that affect the firm, and changes in firm-performance. Examples of event predictors include share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending. Market predictors are predictors that can be constructed using only financial data, such as volume, prices, returns and shares outstanding. Momentum, long-term reversal, and market value of equity are included in our sample of market predictors.

Valuation predictors are ratios, where one of the numbers reflects a market value and the other reflects fundamentals. Examples of valuation predictors include sales-to-price and market-to-book. Finally, fundamental predictors are those that are constructed with financial statement data and nothing else. Leverage, taxes, and accruals are fundamental predictors.

We summarize the in-sample and post-publication returns for the four groups of predictors in Figure 3. The figure shows that all four groups of predictors have post-publication declines in predictability. The post-publication decline is therefore robust; the effect is not driven by a certain type of predictor. The figure shows that market predictors have the largest in-sample mean, averaging 97.5 basis points per month, and the biggest decline post-publication, falling 46.7 basis points to 50.8 basis points per month. Market predictors are constructed using only price and trading data, so they potentially represent violations of the weakest form of market efficiency. The results are consistent with the idea that such predictors are arbitraged more aggressively post-publication.

We more formally test for differences between the four-predictor groups in the regressions reported in Table 4. In each regression monthly returns are regressed on a dummy variable representing one of the four-predictor types. The intercepts in these regressions therefore represent the average returns from the three-predictor types not represented by the dummy variable.

$$R_{it} = \alpha_{i,} + \beta_1 \, Predictor\ Type\ Dummy_i \ + \ e_{it} \quad (2)$$

In the first four regressions we limit the sample to in-sample observations. In the regression in column 1 we include a dummy variable for the Event predictors. The coefficient in this regression is -0.037 and not significant, while the intercept is 0.780 and highly significant. What this shows is that the average monthly return in-sample for the non-Event predictors is 78 basis points per month, and for Event predictors we estimate that the average is 3.7 basis points lower, although the standard error of this estimate is too high for standard statistical significance.

What the effects in the first four regressions of Table 4 show is that, consistent with Figure 3, Market predictors have higher in-sample means. In the second regression the coefficient for the Market predictor dummy is 0.282, although the standard error is too high for the effect to be statistically significant. What we learn from these four regressions is that although Event predictors have higher in-sample means (as shown in Figure 3), the standard error associated with this difference is very large.

Regressions 4-8 in Table 4 test for differences in post-publication declines across the four groups, while controlling for the in-sample mean of each predictor. Table 2 shows that predictors with larger in-sample mean returns have greater post-publication declines, so in these regressions we control for each predictor's in-sample mean return.

$$R_{it} = \alpha_{i,} + \beta_1\, Predictor\ Type\ Dummy_i + In - Sample\ Mean_i + e_{it} \quad (3)$$

The only significant effect in regressions 5-8 is found in regression 6, which has a dummy variable that signifies whether the predictor is a Market predictor. The coefficient is -0.270. What this shows is that after controlling for the in-sample mean of each predictor, Market predictors have post-publication returns that are on average 27 basis points lower than other predictors. It could be the case that because Market predictors require the least amount of information to construct, informed trades are more likely to be attracted to these predictors than the other predictors that require either financial statement information or analyst forecast information to construct.

3.6. Does Costly Arbitrage Play a Role?

Some of the results in the previous tables are consistent with the idea that publication attracts arbitrageurs, which results in lower returns post-publication. As we explain in the Introduction, Pontiff (1996, 2006) and Shleifer and Vishny (1997) point out that costs associated with arbitrage can prevent arbitrageurs from fully eliminating mispricing. By this logic, predictor portfolios consisting more of stocks that are costlier to arbitrage (e.g., smaller stocks, less liquid stocks, stocks with more idiosyncratic risk) should decline less post-publication. If predictor returns are the outcome of rational asset pricing, then we would not expect the post-publication decline to be related to arbitrage costs.

Previous papers in the costly arbitrage literature relate arbitrage costs to differences in returns across stocks within a predictor portfolio (see Pontiff, 2006; Duan, Hu, and McLean, 2010; and McLean, 2010). In contrast, we estimate differences across predictor portfolios. Another difference between our test and the previous literature is that previous studies assume informed traders throughout the entire sample. In this framework, the informed trader had knowledge of the predictor before (and after) the publication date. In contrast, our tests assume that publication provides information to some sophisticated traders which, in turn, causes decay in return-predictability post-publication.

To create the costly arbitrage variables, we perform monthly ranks of all of the stocks in CRSP based on three transaction cost measures: size, bid-ask spreads, dollar volume, and two holding costs measures: idiosyncratic risk and a dividend-payer dummy.

Firm size is measured as the market value of equity. Average monthly spreads are estimated from daily high and low prices using the method of Corwin and Schultz (2012). Dollar volume is the number of shares traded during the past month multiplied by the month-end stock price. Stocks with high dollar volume and low spreads are more liquid, and should therefore be less costly to arbitrage, as should larger stocks.

Idiosyncratic risk limits the amount that an investor will invest in a mispriced stock (Treynor and Black, 1973, and Pontiff, 1996 and 2006). We compute monthly idiosyncratic risk by regressing daily returns on the twelve value-weighted industry portfolios from Ken French's website. For each day, we square that day's residuals and, to correct for autocorrelation, add two times the product of that day's and the previous day's residual. The monthly measure is created by adding up the daily data from a given month.

Pontiff (1996 and 2006) explains that dividends mitigate holding costs since they decrease the effective duration of the position. The intuition is that dividends reduce the future amount of capital devoted to the arbitrage, thus reducing the cumulative holding costs.[10] We use a dummy variable equal to one if a firm paid a dividend and zero otherwise.

Our procedure to estimate the arbitrage cost of each predictor portfolio is a follows. First, for each month, we compute the average cross-sectional ranking for a trait (e.g. size or idiosyncratic risk) among all of the stocks CRSP. Each stock-month observation is therefore assigned a ranking value between 0 and 1. Next, each month, we estimate the average rank for

[10] This result assumes that the level of the mispricing is unaffected by the dividend payout. The result also holds for the case where the level of mispricing is influenced by mispricing, but the relative mispricing is not. For proof, see the appendix in Pontiff (2006).

the stocks that are in either the long or the short sides of each predictor portfolio. This creates a time-series of monthly rank-averages for each trait. We then take the average of each time-series to estimate a single costly arbitrage variable for each predictor. We only use in-sample months to create the costly arbitrage variables, as it could be the case that trading caused by publication has an effect on the costly arbitrage variables.

We report the results from these tests in Table 5. The dependent variable in the regressions reported in Table 5 is a predictor's monthly return. We limit the sample to post-publication observations. The independent variables include the predictor's in-sample mean return, and one of the costly arbitrage variables.

$$R_{it} = \alpha_{i,} + \beta_1\, Arbitrage\ Variable_i + \ In - Sample\ Mean_i + \ e_{it} \quad (4)$$

The results show that there are significantly larger post-publication declines for predictors with lower arbitrage costs. Consistent with costly arbitrage, the coefficients for size, dollar volume, and dividends are all negative, but of these three only the dollar volume coefficient is significant. The dividends coefficient is positive, which is consistent with costly arbitraged, although the coefficient is insignificant. The idiosyncratic coefficient is positive and significant, showing that predictor portfolios consisting of stocks with high idiosyncratic risk decline more post-publication. This result is consistent with Pontiff (2006), who reviews a literature that relates arbitrage costs to alpha across stocks *within* predictor portfolios. This literature finds that return-predictability is stronger in stocks with high idiosyncratic risk, even more so than stocks with high transaction costs.

3.7. Post-Publication Trading Activity in Predictor Portfolios

If academic publication provides market participants with information that they trade on, then this trading activity is likely to affect not only prices, but also other indicators of trading activity. We therefore ask whether turnover, dollar volume, variance, and short interest increase in predictor portfolios during the months after publication. To perform these tests we estimate the regression describe in Equation 1, but replace monthly stock returns with a monthly measure of one of the monthly traits. These traits can time vary for all stocks over the sample period, so we focus on changes in cross-sectional ranks, e.g., we ask whether the variance of stocks in predictor portfolios increase relative to other stocks after the predictor has been published.

Similar to the last section, we compute monthly ranks based on turnover, the dollar value of trading volume, and stock return variance. Turnover is measured as shares traded scaled by shares outstanding, while dollar volume is measured as shares traded multiplied by price. Variance is calculated from monthly stock returns over the preceding thirty-six months. For each predictor portfolio, we compute the average cross-sectional ranking (ranges from 0 to 1) among the stocks that enter either the long or the short side of the characteristic portfolio each month, and test whether the average ranking increases post-publication.

With respect to short interest, we do not compute it's ranking, but instead we subtract the average short interest (shares shorted scaled by shares outstanding) of the long side of each predictor portfolio from the average short interest of the short side of each predictor portfolio. If publication draws short sellers to predictors, then this relative shorting measure should increase post-publication.

We report the results from these tests in Tables 6. The results show that variance and dollar volume are significantly higher during the period that is post-sample but pre-publication, while turnover is not. Hence, there appears to be an increase in trading among predictor portfolio stocks even before a paper is published, suggesting that information from papers may get to some investors before the paper is published. The effects are greatest with dollar volume; comparing the post-sample coeffiicent to the regression intercept shows that the average dollar volume rank of a firm in a predictor portfolio is 2.6% higher out-of-sample but pre-publication as compared to in-sample.

The post-publication coefficients show that variance, turnover, and dollar volume are all significantly higher in predictor portfolios after publication. Comparing the coefificents to the intercept that reflects the average within-sample mean, we see that post-publication the average rank within the characteristic portfolios increases by 1%, 2.2%, and 3% for variance, turnover, and dollar volume respectively.

The final column reports the results from the short interest regression. Recall that the short interest variable is the short interest on the short side minus the short interest on the long side. The coefficients in this regression are reported in percent. If investors recognize that predictor portfolio stocks are mispriced, then there should be more shorting on the short side than on the long side. The intercept is 0.117 (and significant) so the average difference in short interest between the short and long side of the characteristic portfolios is 0.117% before publication. The mean and median levels of short interest in our sample (1976-2012) are 3.45% and 0.77% respectively, so this difference is economically meaningful. This result suggests that some practitioners knew that stocks in the predictor portfolios were mispriced and traded accordingly. This could be because practitioners were trading on the predictor, or it could reflect practitioners trading on other strategies, which happen to be correlated with the predictor. As an example, short sellers might evaluate firms individually with fundamental analyses. The resulting positions might be stocks with low book-to-market ratios, high accruals, high stock returns over the last few years, etc., even though short sellers were not directly choosing stocks based on these traits.

Post-sample, relative shorting increases by 0.187%, and post-publication, relative shorting increases by 0.342%. Economically, the effect represents an increase in relative shorting of three-fold post-publication relative to in-sample (the intercept is 0.117%, which reflects the in-sample mean). So although some practitioners may have known about these strategies before publication, the results here suggest that publication made the effects more widely known.

3.7. The Effects of Publication on Correlations Access Characteristic Portfolios

In this section, we study the effects that publication has on correlations across characteristic portfolios. Simple correlations between predictor portfolios are lower than we expected. The mean pairwise correlation in our study is 0.050 and the median is 0.047. These levels of correlation imply even lower covariance than Green, et al. (2012), who show that R^2 between predictors ranges from 6% to 20%. Our results, and those in Green et al., suggest that multi-characteristic investing is likely to enjoy substantial diversification benefits.

If predictor returns reflect mispricing and if mispricing has common causes (e.g., investor sentiment), then we might expect in-sample predictor portfolios to be correlated with other in-sample predictor portfolios. This effect is suggested in Lee, Shleifer, and Thaler (1991), Barberis and Shleifer (2003), and Barberis, Shleifer and Wurgler (2005). If publication causes arbitrageurs to trade in a predictor, then it could cause a predictor portfolio to become more highly correlated with other published predictors and less correlated with unpublished characteristics.

In Table 7, predictor portfolio returns are regressed on an equal-weighted portfolio of all other predictors that are pre-publication and a second equal-weighted portfolio of all of the other predictors that are post-publication. We include a dummy variable that indicates whether the predictor is post-publication, and interactions between this dummy variable and the pre-publication and post-publication predictor portfolios returns.

The results show that while a predictor is pre-publication, its returns are significantly related to the returns of other pre-publication predictor portfolios. The coefficient or beta for the pre-publication predictor portfolio is 0.704 and and it is statistically significant. In contrast, the beta for a pre-publication portfolio with portfolios that are post-publication is 0.013. These findings are consistent with Lee, Shleifer, and Thaler (1991) and Barberis and Shleifer (2003).

The interactions show that once a predictor is published, its returns are less correlated with the returns of other pre-publication predictor portfolios and more correlated with the returns of other post-publication predictor portfolios. The coefficient for the interaction between the post-publication dummy and the return of the portfolio consisting of in-sample predictors is -0.562 (p-value = 0.047). Hence, once a predictor is published, the beta of its returns with the returns of other yet-to-be-published predictors returns virtually disappears, as the overall coefficient reduces to 0.704 – 0.562 = 0.140. The coefficient for the interaction of the post-publication dummy with the returns of the other post-publication predictors is 0.562 (p-value = 0.063), suggesting that there is a significant relation between the portfolio returns of published predictors and other published predictors.

4. Conclusions

This paper studies 95 predictors that have been shown to explain cross-sectional stock returns in peer reviewed finance, accounting, and economics journals. We compare each predictor's return-predictability over three distinct periods: (i) within the original study's sample period; (ii) outside of the original sample period but before publication; and (iii) post-publication.

We use the period during which a predictor is outside of its original sample but still pre-publication to estimate an upper bound on the effect of statistical biases. We estimate the effect of statistical bias to be about 13%. This is an upper bound, because some investors could learn about a predictor while the study is still a working paper. The average predictor's return declines by 44% post-publication. We attribute this post-publication effect both to statistical biases and to arbitrageurs who observe the finding. Combining this finding with an estimated statistical bias of 13% implies a publication effect of about 31%.

Several of our findings support the idea that cross-sectional predictability is the result of mispricing. First, predictor portfolios with larger in-sample returns and that consist more of stocks that are less costly to arbitrage decline more post-publication. Arbitrageurs should pursue trading strategies with highest after-cost returns, so these results are consistent with the idea that publication attracts sophisticated investors. We further find that variance, turnover, dollar volume, and short interest all increase significantly in predictor portfolios post-publication. This is also consistent with the idea that academic research draws attention to the predictors. Finally, we find that before a predictor is featured in an academic publication, its returns are correlated with the returns of other yet-to-be-published predictors, but its returns are not correlated with those of published predictors This is consistent with behavioral finance models of comovement. After publication, a predictor's correlation with yet-to-be-published predictors is close to zero, and its correlation with already-published predictors becomes significant.

References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, "The cross-section of volatility and expected returns," *Journal of Finance* 61, 259-299.

Amihud, Yakov, 2002, "Illiquidity and stock returns: Cross-section and time-series effects," *Journal of Financial Markets* 5, 31-56.

Amihud, Yakov, and Haim Mendelson, 1986, "Asset pricing and the bid–ask spread," *Journal of Financial Economics* 17, 223–249.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, forthcoming, "Performance of Institutional Trading Desks: An Analysis of Persistence in Trading Costs," *Review of Financial Studies*.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2009, "Value and momentum everywhere," Working paper, New York University.

Barberis, Nicholas, and Andrei Shleifer, 2003, "Style investing," *Journal of Financial Economics* 68, 161-199.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, "Comovement," *Journal of Financial Economics* 75, 283-317.

Bali, Turan G., and Nusret Cakici, 2008, "Idiosyncratic Volatility and the Cross Section of Expected Returns," *Journal of Financial and Quantitative Analysis* 43, 29-58.

Bali, Turan G., Nusret Cakici, and F. Robert Whitelaw, 2011, "Maxing out: Stocks as lotteries and the cross-section of expected returns," *Journal of Financial Economics* 99, 427-446.

Banz, Rolf W., 1981, "The relationship between return and market value of common stocks," *Journal of Financial Economics* 9, 3-18.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, "Comovement," *Journal of Financial Economics* 75, 283-317.

Blume, Marshal E. and Frank, Husic, 1973, "Price, beta, and exchange listing," *Journal of Finance* 28, 283-299.

Bonferroni, C. E. "Il calcolo delle assicurazioni su gruppi di teste." In Studi in Onore del Professore Salvatore Ortu Carboni. Rome: Italy, pp. 13-60, 1935.

Boyer, Brian, 2011, "Style-related Comovement: Fundamentals or Labels?," *Journal of Finance* 66, 307-332.

Brennan, Michael J., 1970, "Taxes, market valuation, and corporate financial policy," *National Tax Journal* 23, 417–427.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong, 2011, "Trends in the cross-section of expected stock returns," Working paper, Emory University.

Cochrane, John H., 1999, "Portfolio Advice for a Multifactor World," *Economic Perspectives* Federal Reserve Bank of Chicago 23, 59-78.

Corwin, Shane A., and Paul Schultz, 2012, "A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices," *Journal of Finance* 67, 719-759.

Drake, Michael S., Lynn Rees, and Edward P. Swanson, 2011, "Should investors follow the prophets or the Bears? Evidence on the use of public information by analysts and short sellers," *Accounting Review* 82, 101-130.

Duan, Ying, Gang Hu, and R. David McLean, 2009, "When is Stock-Picking Likely to be Successful? Evidence from Mutual Funds," *Financial Analysts Journal* 65, 55-65.

Duan, Ying, Gang Hu, and R. David McLean, 2009, "Costly Arbitrage and Idiosyncratic Risk: Evidence from Short Sellers," Journal *of Financial Intermediation* 19, 564-579.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990, "Noise trader risk in financial markets," *Journal of Political Economy* 98, 703–738.

Dichev, Ilia D., 1998, "Is the Risk of Bankruptcy a Systematic Risk?," *The Journal of Finance* 53, 1131-1148.

Dichev, Ilia D., and Joseph D. Piotroski, 2001, "The long-run stock returns following bond ratings changes," *Journal of Finance* 56, 173-203.

Fama, Eugene F., 1976, *Foundations of Finance*, Basic Books, New York (New York).

Fama, Eugene F., 1991, "Efficient capital markets: II," *Journal of Finance* 46, 1575-1617.

Fama, Eugene F., and French, Kenneth R., 1992, "The cross-section of expected stock returns," *Journal of Finance* 47, 427-465.

Fama, Eugene F., and James D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy* 81, 607-636.

Franzoni, Francesco and Jose M. Marin, 2006, "Pension Plan Funding and Stock Market Efficiency," *Journal of Finance* 61, 921-956.

Greenwood, Robin. 2008, "Excess Comovement of Stock Returns: Evidence from Cross-sectional Variation in Nikkei 225 Weights," *Review of Financial Studies* 21, 1153-1186.

Goldstein, Michael, Paul Irvine, Eugene Kandel, and Zvi Weiner, 2009, "Brokerage commissions and Institutional trading patterns," *Review of Financial Studies* 22, 5175-5212

Hanson, Samuel G., and Adi Sunderam, 2011, "The growth and limits of arbitrage: Evidence from short interest," Working paper, Harvard Business School.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2013, "… and the cross-section of expected returns," unpublished working paper, Duke University

Haugen, Robert A, and Nardin L. Baker, 1996, "Commonality in the determinants of expected stock returns," *Journal of Financial Economics* 41, 401-439.

Heckman, James, 1979, "Sample selection bias as a specification error," *Econometrica* 47, 153–161.

Hedges, Larry V., 1992, "Modeling publication selection effects in meta-analysis," *Statistical Science* 7, 246-255.

Hwang, Byoung-Hyoun and Baixiao Liu, 2012, "Which anomalies are more popular? And Why?, Purdue working paper.

Goyal, Amit, and Ivo Welch, 2008, "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies* 21, 1455-1508.

Green, Jeremiah, John R. M. Hand, and X. Frank Zhang, 2012, "The Supraview of Return Predictive Signals," Working paper, Pennsylvania State University.

Grundy, Bruce D. and Spencer J. Martin, J. Spencer, 2001, "Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing," *Review of Financial Studies* 14, 29-78.

Jegadeesh, Narasimhan and Sheridan Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, "Profitability of momentum strategies: An evaluation of alternative explanations," *Journal of Finance* 56, 699-720.

Jegadeesh, Narasimhan and Sheridan Titman, 2011, "Momentum," Working paper, Emory University.

LeBaron, Blake, 2000, "The Stability of Moving Average Technical Trading Rules on the Dow Jones Index," *Derivatives Use, Trading and Regula*tion 5, 324-338.

Leamer, Edward E., 1978, *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, New York: John Wiley & Sons.

Lee, Charles, Andrei Shleifer, and Richard Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75-109.

Lewellen, Johnathan, 2011, "The cross-section of expected returns," Working paper, Tuck School of Business, Dartmouth College.

Lo, Andrew, and Craig MacKinlay, 1990, "Data-snooping biases in tests of financial asset pricing models," *Review of Financial Studies* 3, 431-467.

McLean, R. David, 2010, "Idiosyncratic risk, long-term reversal, and momentum," *Journal of Financial and Quantitative Analysis*, 45, 883-906.

Michaely, Roni, Richard Thaler, and Kent L. Womack, 1995, "Price reactions to dividend initiations and omissions: Overreaction or drift?," *Journal of Finance* 50, 573-608.

Milian, Jonathan A., 2013, "Overreacting to a History of Underreaction?," working paper, Florida International University

Mittoo, Usha, and Rex Thompson, 1990, "Do capital markets learn from financial economists?," working paper, Southern Methodist University

Moskowitz, Tobias, Yao Hua Ooi, and Lasse H. Pedersen, 2010, "Time series momentum," Working paper, New York University.

Muth, John F., 1961, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29, 315–335.

Naranjo, Andy, M. Nimalendran, and Mike Ryngaert, 1998, "Stock returns, dividend yields and taxes," *Journal of Finance* 53, 2029-2057.

Pontiff, Jeffrey, 1996, "Costly arbitrage: Evidence from closed-end funds," *Quarterly Journal of Economics* 111, 1135-1151.

Pontiff, Jeffrey, 2006, "Costly arbitrage and the myth of idiosyncratic risk," *Journal of Accounting and Economics* 42, 35-52.

Ritter, Jay R., 1991, "The long-run performance of initial public offerings," *Journal of Finance* 46, 3-27.

Schwert, G. William, 2003, "Anomalies and market efficiency," Handbook of the Economics of Finance, edited by G.M. Constantinides, M. Harris and R. Stulz, Elsevier Science B.V.

Sharpe, William F., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance* 19, 425-442

Shleifer, Andrei, and Robert W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance* 47, 1343-1366

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

Stein, Jeremy C, 2009, "Presidential address: Sophisticated investors and market efficiency," *Journal of Finance* 64, 1517-1548.

Sullivan, Ryan, Allan Timmermann, and Halbert White, 2001, Dangers of data mining: The cse of calendar effects in stock returns, *Journal of Econometrics* 105, 249-286.

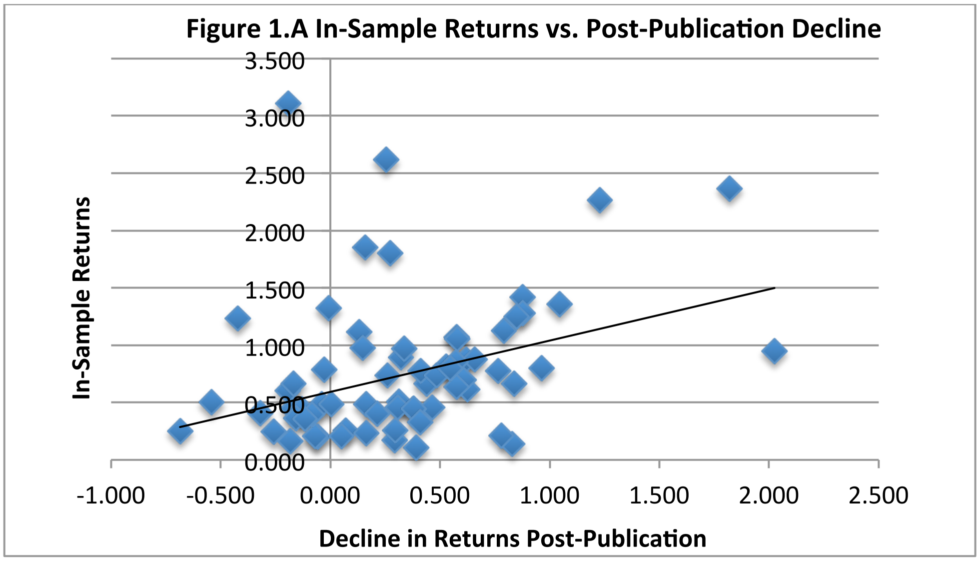
Treynor, Jack, and Fischer Black, 1973, How to Use Security Analysis to Improve Portfolio Selection," *Journal of Business* 46, 66-86.

Vayanos, Dimitri and Paul Woolley, forthcoming, "An institutional theory of momentum and reversal," *Review of Financial Studies*

Wahal, Sunil, and M. Deniz Yavuz, 2009, "Style Investing, Comovement and Return Predictability," Working Paper, Arizona State University.

Figure 1: The relation between in-sample returns and post-publication decline in returns

Figure 1.A plots the relation between in-sample returns and the post-publication decline in returns. For each predictor, we estimate the mean return to a long-short portfolio that contemporaneously buys and sells the extreme quintiles of each predictor characteristic during the sample period of the original study. We then estimate the mean returns for the period after the paper is published through 2012. To be included in the figure, a predictor's in-sample returns had to generate a t-statistic greater than 1.5. 78 of the 95 predictors that we examine met this criterion. The predictor also had to have at least three years of post-publication return data. This excluded 3 of the 78 predictors, resulting in a sample of 76 predictors. Figure 1.B repeats this exercise, only it plots the in-sample t-statistic against the post publication decline.



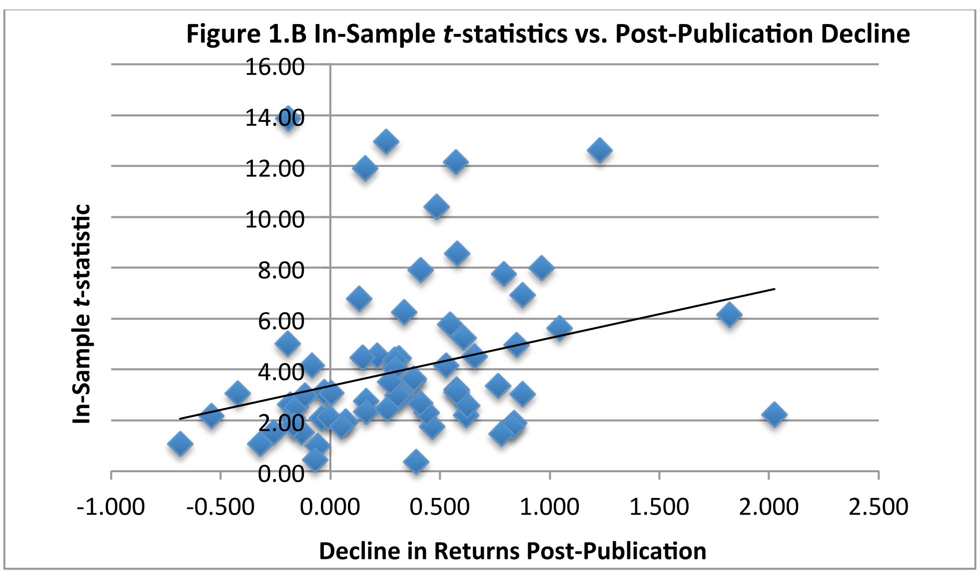


Figure 2: Predictor Return Dynamics around the Sample-End and Publication Dates

Figure 2 explores changes in predictability by examining finer post-sample and post-publication partitions. The figure plots the coefficients from a regression containing dummy variables that signify the last 12 months of the original sample; the first 12 months out-of sample; and the other out-of-sample months. In addition, the publication dummy is split up into six different variables; one dummy for each of the first five years post-publication, and one dummy for all of the months that are at least five years after publication.

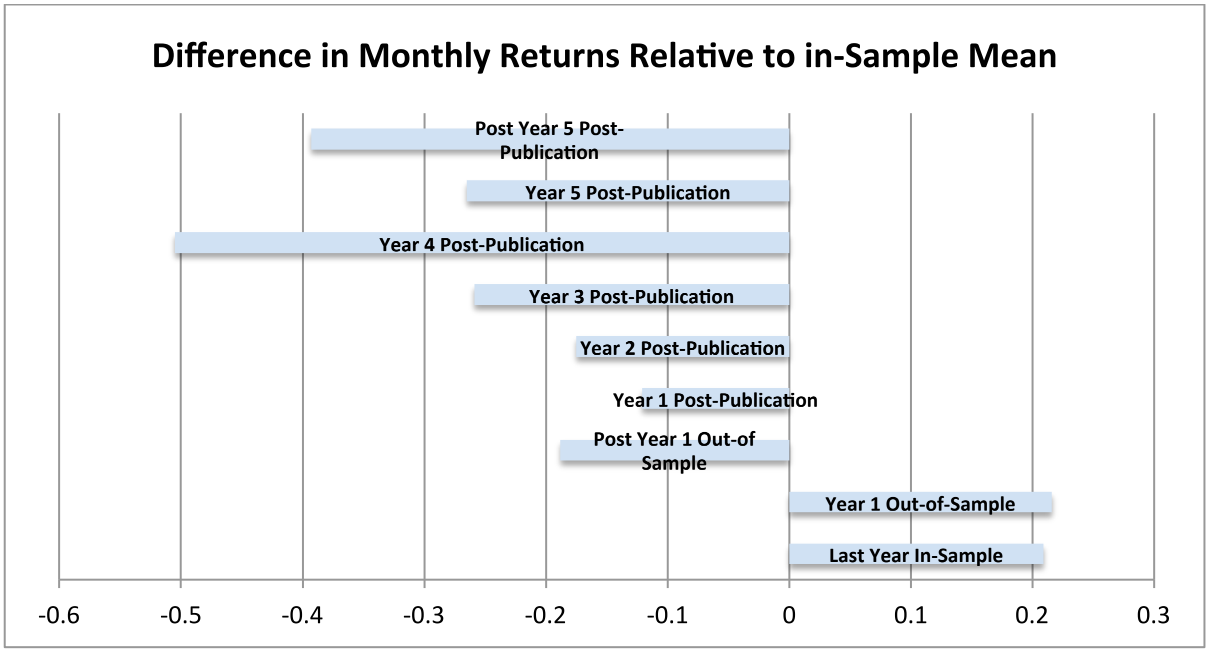


Figure 3: In-sample returns and post-publication declines by predictor type.

This figure graphs the average in-sample returns, post-publication returns, and post-publication decline for four different predictor groups. To conduct this exercise we split our predictors into four groups: (i) Event; (ii) Market (iii) Valuation; and (iv) Fundamentals. Event predictors are those based on corporate events or changes in performance. Examples of event predictors are share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending. Market predictors are predictors that can be constructed using only financial data, such as volume, prices, returns and shares outstanding. Momentum, long-term reversal, and market value of equity (size) are included in our sample of market predictors. Valuation predictors are ratios, where one of the numbers reflects a market value and the other reflects fundamentals. Examples of valuation predictors include sales-to-price and market-to-book. Fundamental predictors are those that are constructed with financial statement data and nothing else. Leverage, taxes, and accruals are fundamental predictors.

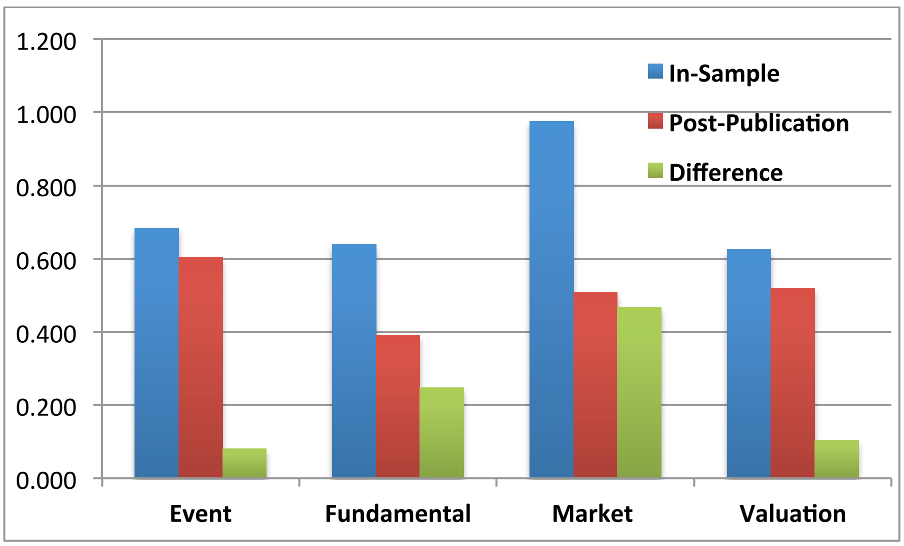


Table 1. Summary Statistics

This table reports summary statistics for the 95 predictors studied in this paper. The mean returns and standard deviations are equal-weighted by predictor. We first estimate the statistic for each predictor, and then taking an equal-weighted average across predictors. Our sample period ends in 2012.

Number of Predictors	95
Predictors with significant returns in-sample:	78 (82%)
Mean Publication Year	2000
Median Publication Year	2001
Predictors from Finance journals	66 (70%)
Predictors from Accounting journals	27 (28%)
Predictors from Economics journals	2 (2%)
Mean Return In-Sample	0.785
Standard Deviation In-Sample	0.562
Mean Observations In-Sample	329
Mean Portfolio Return Out-of Sample	0.685
Portfolio Standard Deviation Out-of-Sample	0.881
Mean Observations Out-of-Sample	55
Mean Return Post-Publication	0.475
Standard Deviation Post-Publication	0.649
Mean Observations Post-Publication	141

Table 2. Regression of predictor portfolio returns on post-sample and post-publication indicators.

The regressions test for changes in returns relative to the predictor's sample-end and publication dates. The dependent variable is the monthly return to a long-short portfolio that is based on the extreme quintiles of each predictor. Post-Sample (S) is equal to 1 if the month is after the sample period used in the original study and zero otherwise. Post-Publication (P) is equal to 1 if the month is after the official publication date and zero otherwise. In-Sample Mean is the mean return of predictor portfolio during the original sample period. t-statistic is the in-sample t-statistic of each predictor portfolio. The regressions are generalized least squares (GLS) regressions that account for covariances across predictors when estimating the regression coefficients. Standard errors are reported in parentheses and clustered on time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. The bottom row reports p-values from tests of whether any declines are 100% of the in-sample mean (the effects disappears entirely), and whether post-sample and post-publication changes in returns are statistically different from one another.

	(1)	(2)	(3)	(4)	(5)
Post-Sample (S)	-0.104	-0.091	-0.086	0.088	0.027
	(0.083)	(0.067)	(0.080)	(0.097)	(0.139)
Post-Publication (P)	-0.351***	-0.300***	-0.320***	-0.070	-0.160
	(0.092)	(0.079)	(0.073)	(0.111)	(0.137)
In-Sample Mean			0.912***		
			(0.044)		
S x Mean				-0.247*	
				(0.137)	
P x Mean				-0.342***	
				(0.131)	
S x t-statistic					-0.029
					(0.018)
P x t-statistic					-0.044***
					(0.016)
Constant	0.795***	0.683***	0.067	0.797***	0.794***
	(0.050)	(0.045)	(0.041)	(0.051)	(0.051)
Predictor FE?	Yes	Yes	No	Yes	Yes
Observations	41,507	49,536	41,507	41,507	41,507
R-squared (Within)	0.001	0.001	0.001	0.001	0.001
Predictors (N)	78	95	78	78	78
Null Hypothesis: S=-1	0.000	0.000	0.000	0.000	0.000
Null Hypothesis: P=-1	0.000	0.000	0.000	0.000	0.000
Null Hypothesis: S=P	0.004	0.004	0.005	NA	NA

Table 3: Time Trend 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 1 if the month is after the sample period used in the original study and zero otherwise. Post-Publication (P) is equal to 1 if the month is after the official publication date and zero otherwise. Time is the number of months divided by 100 post-Jan. 1926. Post-1993 is equal to1 if the year is greater than 1993 and 0 otherwise. All indicator variables are equal to 0 if they are not equal to 1. I-Time is the number of months (in hundreds) after the beginning of the original sample. If the observation falls outside the original sample, I-Time is set to 0. S-Time is the number of months (in hundreds) after the end of the original sample, but before publication. If the observation falls outside this range, S-Time is set to 0. P-Time is the number of months (in hundreds) after the publication date. If the observation is before the publication date, P-Time is set to 0. 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. The regressions are generalized least squares (GLS) regressions that account for covariances across predictors when estimating the regression coefficients. Standard errors are reported in parentheses and clustered on time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table 3: (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time	-0.078***						
	(0.024)						
P-Time		-0.205***		-0.112			
		(0.062)		(0.088)			
S-Time		-0.269		-0.485*			
		(0.183)		(0.282)			
I-Time		0.009		-0.009			
		(0.019)		(0.028)			
1993			-0.111				
			(0.078)				
Post-sample				0.040	-0.116	-0.097	-0.099
				(0.146)	(0.081)	(0.082)	(0.082)
Post Pub.				-0.250*	-0.224*	-0.318***	-0.274***
				(0.150)	(0.118)	(0.087)	(0.095)
1-Month Return						10.410**	
						(4.178)	
12-Month Return							2.044**
							(0.818)
Constant	1.270***	0.743***	0.739***	0.808***	0.737***	0.714***	0.607***
	(0.170)	(0.120)	(0.106)	(0.117)	(0.158)	(0.054)	(0.078)
Observations	41,507	41,507	41,507	41,507	41,507	41,131	40,217
Within R-squared	0.001	0.001	0.001	0.001	0.016	0.007	0.012
Predictor FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE?	No	No	No	No	Yes	No	No

Table 4: Predictor returns across different predictor types

This table tests whether predictor returns and changes in returns post-publication vary across different types of predictors. To conduct this exercise we split our predictors into four groups: (i) Event; (ii) Market (iii) Valuation; and (iv) Fundamentals. Event predictors are those based on corporate events or changes in performance. Examples of event predictors are share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending. Market predictors are predictors that can be constructed using only financial data, such as volume, prices, returns and shares outstanding. Momentum, long-term reversal, and market value of equity (size) are included in our sample of market predictors. Valuation predictors are ratios, where one of the numbers reflects a market value and the other reflects fundamentals. Examples of valuation predictors include sales-to-price and market-to-book. Fundamental predictors are those that are constructed with financial statement data and nothing else. Leverage, taxes, and accruals are fundamental predictors. We regress monthly predictor returns on dummy variables that signify each predictor group. In the regressions 1-4 we limit the sample to in-sample observations. In regressions 5-8 we limit the sample to post-publication observations and include the in-sample mean return of each predictor portfolio as a control variable. The regressions are generalized least squares (GLS) regressions that account for covariances across predictors when estimating the regression coefficients. Standard errors are reported in parentheses and clustered on time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Event	-0.037				0.108			
	(0.134)				(0.070)			
Market		0.282				-0.270*		
		(0.248)				(0.152)		
Valuation			-0.145				0.149	
			(0.167)				(0.183)	
Fundamental				-0.144				0.082
				(0.138)				(0.092)
In-Smpl. Mean					0.789***	0.837***	0.800***	0.799***
					(0.074)	(0.068)	(0.071)	(0.072)
Constant	0.780***	0.693***	0.790***	0.805***	-0.183*	-0.102	-0.182**	-0.173*
	(0.137)	(0.084)	(0.109)	(0.115)	(0.099)	(0.092)	(0.086)	(0.090)
Observations	26,013	26,013	26,013	26,013	11,133	11,133	11,133	11,133
Predictors	78	78	78	78	76	76	76	76

Table 5: Costly arbitrage and the persistence of predictor returns

This regression tests whether arbitrage costs are associated with declines in predictability post-publication. The sample is limited to post-publication months. The dependent variable is a predictor portfolio's monthly long-short return. The independent variables reflect various traits of the stocks in each predictor portfolio. To measure the strength of the traits of the stocks within a portfolio, we do the following. We first rank all of the stocks in CRSP on the trait (e.g., size or turnover), assigning each stock a value between 0 and 1 based on its rank. We then take the average rank of all of the stocks in the portfolio for that month. Finally, we take an average of predictor's monthly trait averages, using all of the months that are in-sample. Hence, in the size regression reported in the first column, the independent variable is the average market value rank of the stocks in the predictor's portfolio during the in-sample period for the predictor. Average monthly spreads are estimated from daily high and low prices using the method of Corwin and Schultz (2012). Dollar volume is shares traded multiplied by stock price. Idiosyncratic risk is daily stock return variance, which is orthogonal to the market and industry portfolios. Dividends is a dummy equal to 1 if the firm paid a dividend during the last year and zero otherwise. The regressions are generalized least squares (GLS) regressions that account for covariances across predictors when estimating the regression coefficients. Standard errors are reported in parentheses and clustered on time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Size	Spreads	Dollar Volume	Idio. Risk	Dividends
In-Sample Mean	0.756***	0.771***	0.731***	0.734***	0.759***
	(0.069)	(0.072)	(0.070)	(0.066)	(0.076)
Characteristic	-0.671	-0.260	-1.638**	1.662***	-0.305
	(0.840)	(0.905)	(0.669)	(0.556)	(0.282)
Constant	0.264	0.007	0.684**	-0.909***	0.068
	(0.428)	(0.421)	(0.317)	(0.302)	(0.178)
Observations	0.000	0.000	0.001	0.000	0.001
R-squared	9,823	9,823	9,823	9,823	9,823

Table 6: Trading activity dynamics in predictor portfolio stocks

This regression models the dynamics of the traits of stocks in predictor portfolios, relative to the predictor's original sample period and publication date. We perform monthly ranks based on turnover, dollar value of trading volume, and stock return variance. Turnover is measured as shares traded scaled by shares outstanding, while dollar volume is measured as shares traded multiplied by price. Variance is calculated from monthly stock returns over the preceding thirty-six months. For each predictor portfolio, we compute the average cross-sectional ranking (ranges from 0 to 1) among the stocks that enter either the long or the short side of the characteristic portfolio each month, and test whether the average ranking increases out-of-sample and post-publication. For short interest (shares shorted scaled by shares outstanding), we take the average short interest in the short quintile for each characteristic, and subtract from it the average short interest in the long quintile. The short interest findings are reported in percent. *Post-sample* is equal to 1 if the month is after the end of the sample, but pre-publication. Post-Sample (S) is equal to 1 if the month is after the sample period used in the original study and zero otherwise. Post-Publication (P) is equal to 1 if the month is after the official publication date and zero otherwise. The regressions are generalized least squares (GLS) regressions. The regressions are generalized least squares (GLS) regressions that account for covariances across predictors when estimating the regression coefficients. Standard errors are reported in parentheses and clustered on time. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Variance Rank	Trading Volume Rank	Dollar Volume Rank	Difference in Short Interest
Post-Sample (S)	0.007**	0.007	0.012*	0.187**
	(0.003)	(0.007)	(0.007)	(0.085)
Post-Publication (P)	0.005*	0.011*	0.014**	0.342***
	(0.003)	(0.006)	(0.005)	(0.128)
Constant	0.515***	0.497***	0.469***	0.117**
	(0.001)	(0.002)	(0.002)	(0.022)
Observations	42,423	42,400	42,328	32,298
R-squared	0.006	0.005	0.009	0.024
Predictor FE?	Yes	Yes	Yes	Yes
Null: S=P	0.10	0.17	0.07	0.09

Table 7: Regressions of predictor returns on return indices of other predictors

This regression models the returns of each predictor relative to the returns of other predictors. The dependent variable is a predictor's monthly long-short return. Post-Publication (P) is equal to 1 if the month is after the official publication date and zero otherwise. Mean In-Sample is the equal-weighted return of all other unpublished predictor portfolios. Mean Post-Publication is an equal-weighted return of all other published predictor portfolios. The regressions are generalized least squares (GLS) regressions. Standard errors are clustered on time and reported in parentheses. P-values are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Return
Mean In-Sample	0.704***
	(0.032)
Mean Post Pub.	0.013*
	(0.007)
P x Mean In-Sample	-0.562***
	(0.047)
P x Mean Post-Pub.	0.562***
	(0.063)
Post-Publication (P)	-0.140***
	(0.054)
Constant	0.229***
	(0.029)
Observations	34,293
Predictors	78
Within R-squared	0.038