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

June 30, 2014

Abstract

We study the out-of-sample and post-publication return-predictability of 95 characteristics that published academic studies show to predict cross-sectional stock returns. We estimate an upper bound decline in predictability due to statistical bias of 25%, and a post-publication decline, which we attribute to both statistical bias and informed trading, of 56%. Our findings support the contention that investors learn about mispricing from publications. Post-publication declines are greater for predictors with larger in-sample returns, and returns are lower for predictors concentrated in stocks with low idiosyncratic risk and high liquidity. Post-publication, predictor portfolios exhibit 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, Ivey Business School, Pontificia Universidad Catholica de Chile , 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, Cam Harvey, Clifford Holderness, Darren Kisgen, Borja Larrain, Owen Lamont, Jay Ritter, Andrei Shliefer, Paul Schultz, Bruno Skolnik, Jeremy Stein, Ken Singleton, Noah Stoffman, Matti Suominen, Allan Timmermann, Michela Verado, Artie Woodgate, Jianfeng Yu, three anonymous, and an anonymous associate editor 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 out-of-sample cross-sectional predictability. For an analysis of the performance of out-of-sample time-series predictability, see LeBaron (2000) and Goyal and Welch (2008). For an analysis of cross-sectional predictability using international data, see Fama and French (1998), Rouwenhorst (1998), and McLean, Watanabe, and Pontiff (2009). 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 solely 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: specification selection bias, sample selection bias, and multiple testing bias. Leamer (1978) points out that a bias arises when the choice of a method is influenced by the method's result. Lo and MacKinlay (1990) study a version of the specification selection bias in finance, and refer to it as the "data snooping bias." The sample selection bias is studied in Heckman (1979); this bias arises when the sample construction is influenced by the result of the test.[3] A multiple testing bias arises when researchers conduct multiple tests of the same hypothesis. This bias 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 due to the growth of finance research. 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. Publication dates are clear and reflect an important milestone for the dissemination of research. Knowledge about a finding is more widespread after publication. Despite this, the publication date is unlikely to be associated with abrupt changes in return-predictability. On the one hand, papers are often presented at conferences and distributed

[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.

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. Previous versions of this paper showed that the consideration of different dates around the publication period had little impact on the results. For example, we generated similar regression results using SSRN posting dates.[4]

Whether return predictability changes post-publication or not depends on what causes the predictability. Cochrane (1999) explains that if predictability reflects risk it is likely to persist: "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 entirely reflects rational expectations, then publication will not convey information that causes a rational agent to behave differently. Thus, pre- and post-publication return-predictability should be similar.

Alternatively, if publication draws the attention of sophisticated investors who learn about mispricing (or risk-reward "deals") and these investors trade against the mispricing, then we expect the returns associated with a predictor to disappear or at least decay after the paper is published.[5] Decay, as opposed to disappearance, is likely to occur if impediments prevent arbitrage from fully eliminating mispricing. Examples of such impediments include systematic noise trader risk (Delong, Shleifer, Summers, and Waldman (1990)) and idiosyncratic risk 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.

[5] We do not distinguish between mispricing and risk-reward deals since both are inconsistent with rational expectations. Liu, Lu, Sun, and Yan (2014) develop a model of risk-reward deals and learning that is framework for our findings.

transaction costs, Pontiff (1996, 2006). These effects can be worsened by the principal-agent relations that exist between investors and investment professionals, Shleifer and Vishny (1997)).[6]

Findings. We conduct our analysis using 95 different characteristics from 78 different studies. Using long-short portfolio strategies that buy and sell extreme quintiles that are based on each predictor, we confirm significant returns for 80 of the characteristics in-sample; for 15 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 80 characteristics with significant in-sample returns. As we mention above, the post-sample but pre-publication period is useful for estimating statistical bias. The average in-sample monthly return is 66.7 basis points and the average post-sample but pre-publication monthly return is 50.2 basis points, so, on average, return-predictability declines by 25% during this period. This finding is marginally significant. Our 25% estimate is probably too high, since some traders are likely to learn about the predictor before publication and their actions will cause some decay that is captured in the 25%.

The average, in-sample return of 66.7 basis points per month, decays to 29.3 basis points per month out-of-sample. This is a 56% post-publication decay. Combining this finding with an estimated statistical bias of 25% implies a lower bound on the publication effect of about 31%. We can reject the hypothesis that return-predictability disappears entirely, and we can reject the hypothesis that post-publication return-predictability does not change. The post-publication decline is robust to a general time trend, to time indicators used by other authors, and to time fixed effects.

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

The decay in portfolio returns is larger for predictor portfolios 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 likely to represent violations of weak form market efficiency. Post-publication returns are lower for predictors that are less costly to arbitrage; i.e., predictor portfolios with liquid stocks and low idiosyncratic risk stocks. Our findings are consistent with mispricing accounting for some or all of the original return predictability, and investors learning about mispricing.

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 0.046. This finding is in-line with Green, Hand, and Zhang (2013) who find a 0.09 correlation among 60 quantitative portfolios. 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

[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.

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

Our findings are related to contemporaneous research that investigates how the magnitude of sophisticated capital affects anomaly returns (Hanson and Sundareram, 2014, Kokkonen and Suominen 2014, and Akbas, Armstrong, Sorescu, and Subrahmanyam, 2014). Unlike these papers, we do not consider proxies for variation in sophisticated capital levels. Rather, we investigate the discrete role of academic publications in transmitting information to sophisticated investors.

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. Returns are equally weighted unless the primary study presents value-weighted portfolio results (e.g., Ang, Hodrick, Xing, and Zhang, 2006).

We estimate each predictor's return-predictability by computing the return of a portfolio that each month invests in stocks in the top 20th percentile of the characteristic (the strategy's long-side) minus the return of a portfolio that invests in stocks in the bottom 20th percentile of the characteristic. 16 of our 95 predictors are indicator variables. For these cases, if the indicator variable that is associated with returns that outperform the market, the long-short portfolio return formed by buying the indicated stocks and shorting an equal weighted portfolio of all other stocks. For indicators that are associated with underperformance, a similar portfolio is formed using the all other stocks as the long side, and the indicated stocks as the short side.

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 demarcation for estimating statistical bias. The publication date, however, provides only a proxy for when market participants learning about a predictor. 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 80 (95 – 15) or 84% of the predictor's produce significant in-sample returns and are used in the paper's primary tests.

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 83 or 87% of them. In cases where we fail to find in-sample predictability, we are usually able to reconcile our attempts with the original paper. For example, in some cases, the original paper demonstrates abnormal returns from an event study, and these effects don't survive in monthly cross-sectional regressions. In other cases, we do not have exactly the same data used by the original authors.

Admittedly, the decision to use a t-statistic cut-off of 1.50 is arbitrary. The decision is 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.

We define the publication date as the year and month of the journal's issue. For this date convention, the average length of time between the end of the sample and publication is 55

[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.

months. For comparison, the average original in-sample span is 327 months, and the average post-publication span is 151 months. Our sample ends in 2012. 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.

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):

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

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 predictor portfolios are low, averaging only 0.049. Estimation is conducted with a panel regression with predictor portfolio fixed-effects. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. Other specifications were considered (e.g., clustering on time) with similar results.

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 source 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 much of the return-predictability in academic studies is the outcome of data-mining.

Instead, if predictors' returns are entirely 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 80 predictors for which we found significant in sample predictability. The post-sample coefficient in this regression is -0.165 and statistically significant at the 10% level. This means that our best estimate of the post-sample decline 16.5 basis points. The post-publication coefficient is -0.374, and it is significant at the 1% level. This means that on average predictor portfolios are 37.4 basis points lower post-publication compared to before publication. Table 1 shows that the average predictor has an in-sample mean return of 66.7 basis points per month. Hence, the post-sample and post-publication declines relative to the in-sample mean are 25% and 56% respectively.

The regression in the second column includes all 95 predictors, and therefore includes the 15 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 -0.125 and -0.318 respectively in column 2, similar to the results in column 1, although the post-sample coefficient is no longer significant. The average in-sample return for all 95 predictors is 0.571 (not in tables), so the post-publication decay in percentage terms is similar if these other 15 predictors are included. The average return in-sample is smaller because we are including 15 predictors that do not have significant in-sample predictability.

In the regression reported in third column we exclude the predictor fixed effects 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 13.5 basis points out-of-sample, and then a significant decline of 34.5 basis points post-publication.

At the bottom of Table 2, 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 post-publication return-predictability is significantly higher than the decline in out-of-sample, but pre-publication return-predictability. 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 revisits predictor fixed effects. 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 fixed-effects.

In column 4 the coefficient for post-sample is 0.172 and insignificant, while the coefficient for the post-sample interaction with the in-sample mean is -0.497. As we mention above, the mean predictor has an in-sample mean return of 0.667 (see Table 1), so the overall post-sample effect is 0.172 + (-0.497 x 0.667) = -0.159, similar to the post-sample coefficient in column 1. The in-sample mean has a standard deviation of 0.418. (see Table 1) Hence, a one standard deviation increase in the in-sample mean leads to an additional -0.497 x 0.418 =0.207 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 0.012 and insignificant, while the post-publication interaction is -0.525 and highly significant. The average predictor therefore has a post-publication decline of 0.012 + (-0.525 x 0.667) = -0.338, which is similar to the effect estimated in column 1. A one standard deviation increase in the in-sample mean leads to an additional -0.525 x 0.418 = -0.219 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 0.063 and insignificant, while the post-sample t-statistic interaction coefficient is -0.056 and highly significant. The in-sample t-statistics have a mean of 3.94 and a standard deviation of 2.49 (not reported in tables). Hence, the regression estimates an average decline of 15.8 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 13.9 basis points.

The post-publication coefficient in column 5 is -0.104 and insignificant, while the post-publication interaction coefficient is -0.068. Hence, the average predictor portfolio has a post-publication decline of about 37.2 basis points. A one standard deviation increase in the in-sample t-statistic leads to an additional 16.9 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. Some caution is needed in interpreting this figure. Although the estimates in this figure are interesting, statistical power is lost from partitioning the results, and theory does not guide us regarding the appropriate partitions.

The publication process often takes years. This gives unscrupulous 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 virtually zero, showing that the first 12 months post-sample has on average the same returns as compared to 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 20 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 33.7 basis points; in the fourth year it is 62.3 basis points; and in the fifth year it is 19.1 basis points. After the fifth year predictors' returns are on average 41.7 basis points lower as compared to in-sample.

Some readers suggest that we examine post-publication returns as a function of the persistence of the predictor. Initially, decay may be muted if new capital flows into portfolios that are determined by a persistent predictor. For example, new flows into high book-to-market stocks might cause a temporary increase in the returns of book-to-market portfolios. This would not occur in portfolios that are formed on less persistent predictors, such as last month's stock return. In an earlier version of the paper, this possibility was considered. We found some evidence that portfolio returns to more persistent predictors decayed least following publication, but our test lacked statistical power to reject the null.

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, which 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 possibilities 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 predictor fixed effects. 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 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.

In column three, we relate decay to a time trend, the post-1993 indicator, and the post-sample and post-publication indicator variables. This specification causes the coefficient on the time trend to become more negative. The standard error of this coefficient increases, although the null is still comfortably rejected at the 1% level. The post-1993 dummy variable is now *positive* and statistically significant. The post-publication coefficient is very similar as in table 2, -0.346, and statistically significant. Thus, consideration of a time trend and a 1993 break has little impact on post-publication return decay.

In Column 4, we consider time trends within each subperiod, while still estimating discrete post-sample and post-publication declines. We construct three time trend variables. *I-Time* is the number of months (divided by 100) since the sample started. After the in-sample period ends, *I-Time* retains the value of the last month within the in-sample period. *S-Time* is zero before the predictor's sample period ends. During the post-sample, pre-publication period *S-Time*'s value is the number of months (divided by 100) since the original sample ended. Post-publication, *S-Time* retains the value of the last month before publication. Finally, *P-Time* is zero before the publication date. After the publication date, *P-Time*'s value is the number of months (divided by 100) since the publication date.

The results reported in column 4 show that this specification produces negative but insignificant coefficients on all of the time variables but *S-Time*, which is positive but insignificant. The regressors are correlated, and such multicollinearity increases the regression standard errors, thereby reducing statistical significance. Economically, the coefficients suggest decay in predictability throughout the life of a predictor, and that the decay is more pronounced post-publication. Taken together with our other specifications, the tests here suggest that there is both a discrete drop in returns around the publication date, along with a downward trend in returns that goes on for some time post-publication. These results are consistent with the patterns in Figure 2.

Another way to control for time effects while studying the effect of publication is to include time fixed effects. Time fixed effects are highly correlated with our publication indicators; a (unreported) regression of the publication indicators on time fixed effects yields an R^2 of 0.46, so including time fixed effects removes about half of the variation in the publication indicators. We report results from a specification that includes time fixed effects in column 5. This regression estimates declines of 16.2 basis points out-of-sample, and 26.6 basis points post-publication; both coefficients are significant at the 10% level. Based on the average in-sample return of 66.7 basis points, this specification implies a sizeable 39.9% drop in post-publication predictability, and this is after all time effects have been removed.

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 post-publication coefficient remains significant in each of these regressions, suggesting a post-publication decline of about 30 basis points once past returns are considered.

Previous versions of the paper considered whether or not decay was related to the cumulative number of academic citations generated by publication that introduced the portfolio returns associated with the predictor. Once we controlled for publication date, this measure had little incremental value in explaining decay.

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 98.2 basis points per month, and the biggest decline post-publication, falling 45.9 basis points to 52.3 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, a post-publication dummy, and an interaction between the post-publication and predictor type variables.

$$R_{i,t} = \alpha_{i,} + \beta_1\, Post\ Publication\ Dummy_i + \beta_2\, Predictor\ Type\ Dummy_i\ +$$

$$\beta_3\, Post\ Publication\ Dummy_i\ x\ Predictor\ Type\ Dummy_i + \ e_{it} \quad (2)$$

The coefficients for the non-interacted variables in Table 4 test whether the in-sample average returns of a group are different than those of the other groups. The results show that, compared to the other categories of predictors, market-based predictors have significantly higher in-sample returns, while event and fundamental predictors have lower in-sample returns.

The coefficients for the interactions test whether post-publication declines vary across the groups. None of the interactions are significant, although economically the decline for market-based predictors is largest. Moreover, we can estimate differences in expected returns post-publication by adding the type coefficient to the interaction coefficient (β_2 + β_3). These results are reported in the bottom two rows of Table 4. The results show that expected returns post-publication are not significantly different across the four groups. Hence, although in-sample returns are higher for market-based predictors, post-publication this type of predictor does not have higher expected returns.

3.6. Does Costly Arbitrage Play a Role?

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 the post-publication decline should not 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. We also create a costly arbitrage index, which is the first principal component of the five costly arbitrage variables.

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. Large stocks, stocks with high dollar volume and stocks with low spreads are more liquid, and should therefore be less costly to arbitrage. Hence, we expect returns to be lower in predictor portfolios concentrated in such stocks.

Idiosyncratic risk limits the amount that an investor will invest in a mispriced stock (Treynor and Black, 1973, and Pontiff, 1996 and 2006), so we expect returns to be higher in predictor portfolios concentrated in high idiosyncratic risk stocks. We compute monthly idiosyncratic risk by regressing daily returns on the twelve value-weighted industry portfolios from Ken French's website. We estimate a regression for each stock using the last 24 months of daily data. 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 idiosyncratic risk measure is created by adding up the daily sum of residual products from a given month. If the industry factor model regression contains less than 30 observations, the stock is not assigned an idiosyncratic risk measure for that 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.[9] We use a dummy variable equal to one if a firm paid a dividend and zero otherwise. We expect returns to be lower in predictor portfolios concentrated in stocks that pay dividends.

The costly arbitrage index is based on the first principal component of the five costly arbitrage variables. A higher value if the index is associated with lower arbitrage costs, and therefore lower expected portfolio returns. The index has positive correlations with the size, dividends, and dollar volume variables, and negative correlations with the spreads and idiosyncratic risk variables.

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 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 estimate the following regression equation:

$$R_{i,t} = \alpha_{i,} + \beta_1\, Post\ Publication\ Dummy_{i,t} + \beta_2\, Arbitrage\ Cost_i\ +$$
$$\beta_3\, Post\ Publication\ Dummy_{i,t}\ x\ Arbitrage\ Cost_i + \ e_{it}\ (3)$$

[9] 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).

Costly arbitrage predicts that post-publication, predictors that are more difficult to arbitrage should have higher expected returns. Hence, the sum of the costly arbitrage coefficient (β_2) plus the coefficient for the interaction between the post-publication dummy and the arbitrage cost variable (β_3) should reflect higher expected returns for predictors that are more costly to arbitrage. The sum of these coefficients ($\beta_2 + \beta_3$) should therefore be negative for the size, dollar volume, and dividends regressions, and positive for the spreads and idiosyncratic risk regressions.

Table 5 largely supports the notion that some sophisticated traders exert price pressure pre-publication, but the price pressure is tempered by arbitrage costs. If some sophisticated traders implement predictor strategies pre-publication, we expect portfolios with higher arbitrage costs to have higher returns. This effect is ascertained from the slopes on the non-interacted arbitrage cost variables (β_2). Five of the six specifications have slopes with the expected sign, and four of these are statistically significant. The dividend coefficient has the expected sign, but it is insignificant. Dollar volume produces a slope in the opposite direction—predictor portfolios concentrated in stocks with high dollar volume of trading tend to have higher in-sample returns, however the coefficient is insignificant.

Post-publication knowledge of a predictor should be widespread, and we therefore expect portfolios that are easier to arbitrage to have lower post-publication returns. As we explain above, the level of returns is measured by adding the non-interacted slope with interacted slope. All six of these sums have the expected sign, and five of the six are statistically significant. Only the sum of the dividend coefficients is insignificant.

For brevity, we do not report a specification that includes, simultaneously, all five of the primary costly arbitrage variables and all five of the interactions. Caution is needed in interpreting such results due to high correlation between right-hand-side variables. Regarding in-sample returns, idiosyncratic risk is the only costly arbitrage variable that commands a statistically significant slope with the expected sign. Post-publication, returns are higher for predictor portfolios that contain stocks with more idiosyncratic risk. The post-publication effects for spreads and dollar volume have the expected sign, but are not significant. Idiosyncratic risk's overall p-value is 0.001. This finding is consistent with Pontiff's (2006) review of the literature that leads him to conclude, "idiosyncratic risk is the single largest cost faced by arbitrageurs."

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 should affect not only prices, but also other indicators of trading activity. We therefore test 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 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 a predictor portfolio increases relative to other stocks in the same cross-section, 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 rankings, 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 Table 6. The results show that variance, trading volume, and dollar volume are significantly higher during the period that is post-sample but pre-publication. 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 coefficient to the regression intercept shows that the average dollar volume rank of a firm in a predictor portfolio is 2.2% 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 coefficients 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%, 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 0.173 (and significant) so the average difference in short interest between the short and long side of the characteristic portfolios is 0.173% 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 0.125%, and post-publication, relative shorting increases by 0.354%. Economically, the effect represents an increase in relative shorting of two-fold post-publication relative to in-sample (the intercept is 0.173%, 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. These short interest results are consistent with Hanson and Sunderam (2014), who use short interest as a proxy for sophisticated investors, and find that increases in short interest are associated with lower future returns in value and momentum stocks.

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 0.049 and the median is 0.047. These levels of correlation imply even lower covariance than Green, et al. (2012), who show that R^2's between predictor return portfolios 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 returns of 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 before publication predictor returns are significantly related to the returns of other pre-publication predictor portfolios. The coefficient or beta for the pre-publication predictor portfolio is 0.758 and it is statistically significant. In contrast, the beta for a pre-publication portfolio with portfolios that are post-publication is -0.001 and insignificant. 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 -0.706 and highly significant. 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 0.758 – 0.706 = 0.052. The coefficient for the interaction of the post-publication dummy with the returns of the other post-publication predictors is 0.626 and significant at the 1% level, 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. Forming portfolios based on the extreme quintiles for each predictor, we find significant in-sample return predictability for 80 of the 95 predictors. For these 80 predictors, 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 25%. 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 56% post-publication. We attribute this post-publication effect both to statistical biases and to the price impact of sophisticated traders. Combining this finding with an estimated statistical bias of 25% implies a publication effect of 31%.

Several of our findings support the idea that a portion or all of the original cross-sectional predictability is mispricing. First, predictor portfolios with larger in-sample returns decline more post-publication, and strategies concentrated in stocks that are more costly to arbitrage have higher expected returns post-publication. Arbitrageurs should pursue trading strategies with the 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 especially short interest 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

Akbas, Ferhat, Will J. Armstrong, Sorin Sorescu, and Aanidhar Subrahmanyam, 2014, "Time Varying Market Efficiency in the Cross-Section of Expected Stock Returns," Unpublished working paper UCLA.

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.

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 French, Kenneth R., 1998, "Value versus Growth: The International Evidence," *Journal of Finance* 53, 1975-1999.

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, 2014, "The growth and limits of arbitrage: Evidence from short interest," *Review of Financial Studies* 27, 1238-1286.

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.

Liu, Qi, Lei Lu, Bo Sun, and Hongjun Yan, 2014, A model of anomaly discovery, Working paper, Yale School of Management.

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.

McLean, David, Jeffrey Pontiff, and Akiko Watanabe, 2009, "Share Issuance and Cross-Sectional Returns: International Evidence," *Journal of Financial Economics* 94, 1-17.

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.

Rouwenhorst, K. Geert, 1998, International Momentum Strategies, *Journal of Finance* 53, 267-284.

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 case of calendar effects in stock returns, *Journal of Econometrics* 105, 249-286.

Suominen, Matti and Joni Kokkonen, 2014, "Hedge Funds and Stock Market Efficiency," Unpublished working paper, Aalto University.

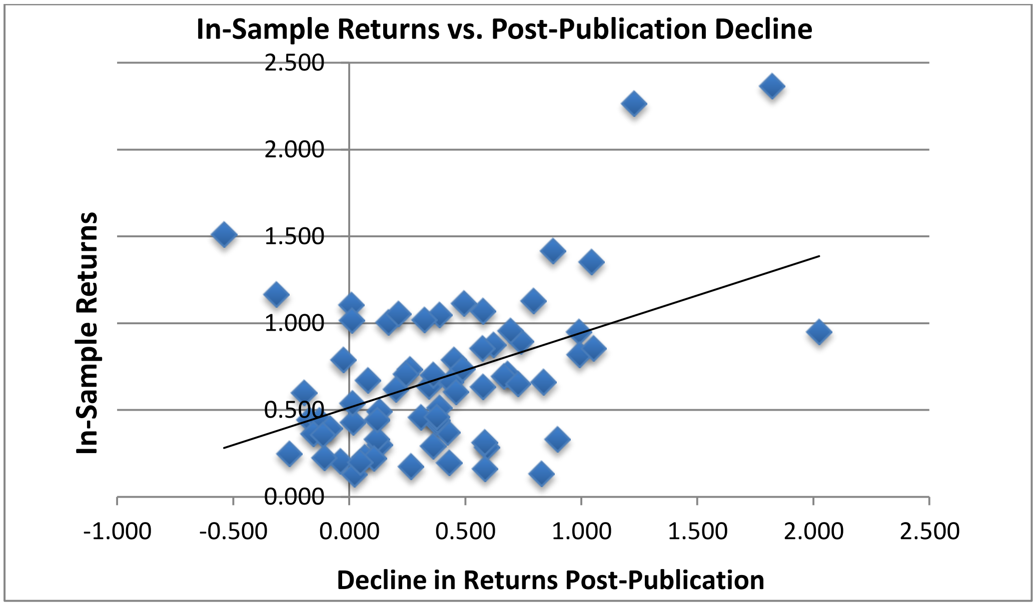
Treynor, Jack, and Fischer Black, 1973, How to Use Security Analysis to Improve Portfolio Selection," *Journal of Business* 46, 66-86.

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. 80 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 10 of the 80 predictors, resulting in a sample of 70 predictors. Figure 1.B repeats this exercise, only it plots the in-sample t-statistic against the post publication decline



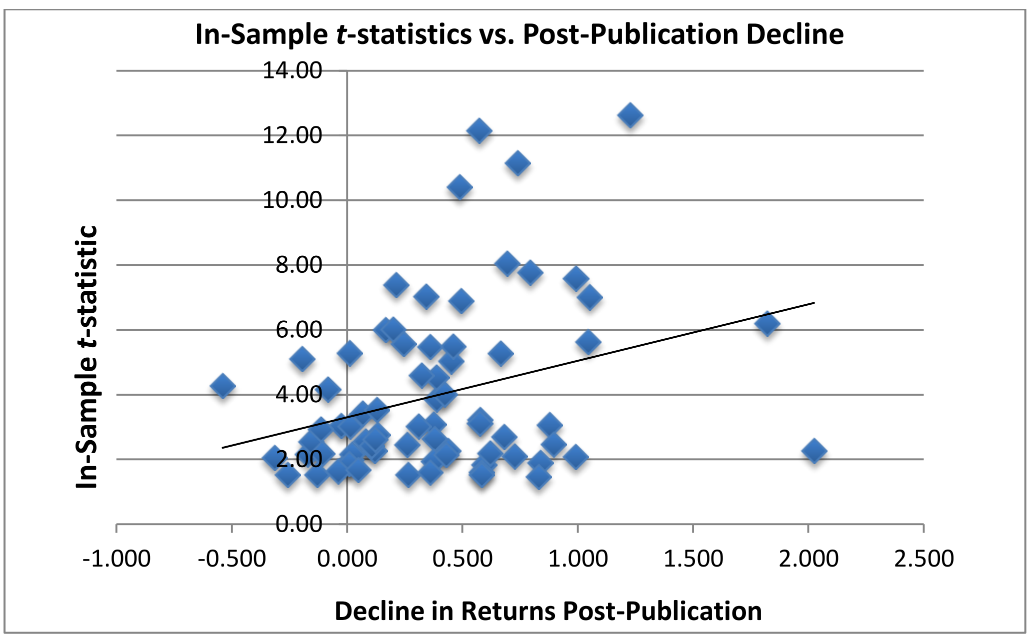


Figure 2: Predictor Return Dynamics around the Sample-End and Publication Dates

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.

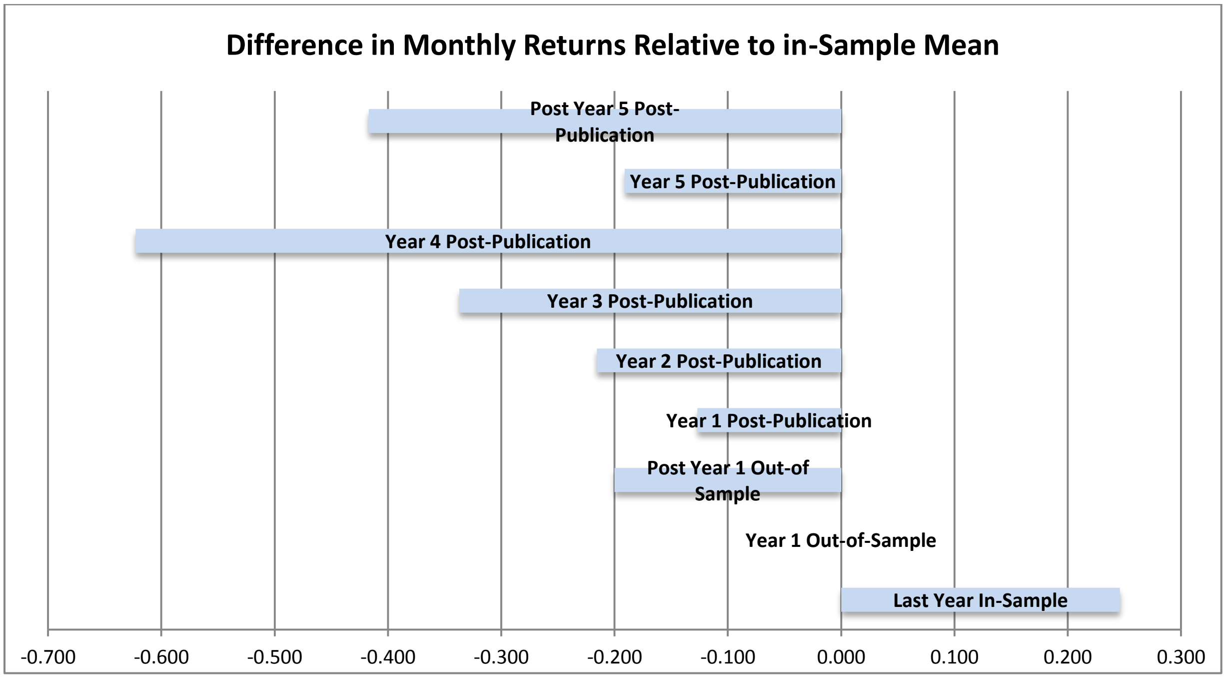


Figure 3: In-sample returns and post-publication declines by predictor type.

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.

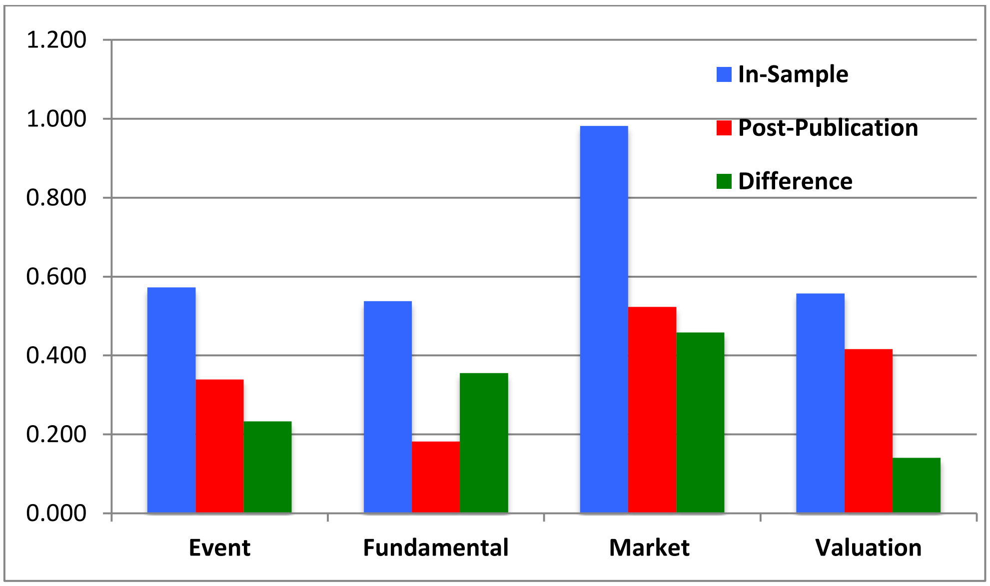


Table 1. Summary Statistics

This table reports summary statistics for the predictors studied in this paper. The returns are equal-weighted by predictor. The standard deviations are the standard deviations of returns across predictors. We only measure returns for the 80 predictors with significant in-sample returns. We first estimate the statistic for each predictor, and then take an equal-weighted average across predictors. The predictor also had to have at least three years of post-publication return data to be included for the estimate of the post-publication returns. This excluded 8 of the 78 predictors, resulting in a sample of 70 predictors. Our sample period ends in 2012.

Number of Predictors	95
Predictors with significant returns in-sample:	80 (84%)
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 Portfolio Return In-Sample	0.667
Standard Deviation In-Sample	0.418
Mean Observations In-Sample	327
Mean Portfolio Return Out-of Sample	0.464
Standard Deviation Out-of-Sample	0.721
Mean Observations Out-of-Sample	55
Mean Portfolio Return Post-Publication	0.294
Standard Deviation Post-Publication	0.499
Mean Observations Post-Publication	151

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. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. 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.

Variables	(1)	(2)	(3)	(4)	(5)
Post-Sample (S)	-0.165*	-0.125	-0.135	0.172	0.063
	(0.094)	(0.085)	(0.089)	(0.130)	(0.156)
Post-Publication (P)	-0.374***	-0.318***	-0.345***	0.012	-0.104
	(0.112)	(0.100)	(0.100)	(0.125)	(0.162)
In-Sample Mean			0.816***		
			(0.075)		
S x Mean				-0.497**	
				(0.249)	
P x Mean				-0.525***	
				(0.197)	
S x t-statistic					-0.056**
					(0.030)
P x t-statistic					-0.068***
					(0.025)
Predictor FE?	Yes	Yes	No	Yes	Yes
Observations	41,530	49,701	41,530	41,530	41,530
R-squared	0.001	0.011	0.010	0.001	0.010
Predictors (N)	80	95	80	80	80
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	0.035	0.028	0.028	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. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table 3: (Continued)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time	-0.082*** (0.022)		-0.095*** (0.034)				
I-Time				-0.041 (0.029)			
S-Time				0.236 (0.173)			
P-Time				-0.102 (0.104)			
1993		-0.126 (0.088)	0.343** (0.141)				
Post-sample			-0.179* (0.097)	-0.163 (0.118)	-0.162* (0.098)	-0.143 (0.094)	-0.127 (0.097)
Post Pub.			-0.346** (0.146)	-0.335 (0.208)	-0.266* (0.146)	-0.333*** (0.112)	-0.278*** (0.116)
1-Month Return						0.100*** (0.017)	
12-Month Return							0.020*** (0.004)
Observations	41,530	41,530	41,530	41,507	41,530	41,530	40, 530
R-squared	0.001	0.001	0.010	0.015	0.120	0.019	0.016
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. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

VARIABLES	(1)	(2)	(3)	(4)
Post-Publication (P)	-0.233***	-0.316***	-0.320***	-0.296***
	(0.068)	(0.121)	(0.104)	(0.114)
P x Market	-0.195			
	(0.205)			
Market	0.407***			
	(0.098)			
P x Event		0.087		
		(0.116)		
Event		-0.148**		
		(0.063)		
P x Valuation			0.188	
			(0.194)	
Valuation			-0.146**	
			(0.091)	
P x Fundamental				-0.037
				(0.159)
Fundamental				-0.210***
				(0.054)
Constant	0.545***	0.716***	0.694***	0.726***
	(0.045)	(0.064)	(0.057)	(0.059)
Observations	41,530	41,530	41,530	41,530
R-squared	0.003	0.001	0.001	0.001
Predictors	80	80	80	80
Type + (P x Type)	0.212	-0.061	0.042	-0.247
p-value	0.212	0.482	0.396	0.082

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 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 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. Index is the first principal component of the other five measures. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. The bottom two rows test whether the sum of the costly arbitrage variable (CA) plus the interaction between the costly arbitrage variable and publication (P x CA) is statistically different than zero.

Table 5 (Continued)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Post-Pub. (P)	0.098	-0.436	0.581**	-0.376	-0.290	-0.272***
	(0.411)	(0.327)	(0.285)	(0.590)	(0.299)	(0.095)
P x Size	-0.695					
	(0.726)					
Size	-0.821**					
	(0.403)					
P x Spreads		0.385				
		(0.830)				
Spreads		0.815**				
		(0.386)				
P x Dol.Vol.			-1.853***			
			(0.625)			
Dollar Volume			0.556			
			(0.405)			
P x Idio. Risk				0.218		
				(1.412)		
Idio. Risk				2.691***		
				(0.523)		
P x Dividends					0.015	
					(0.497)	
Dividends					-0.228	
					(0.214)	
P x Index						-0.036
						(0.032)
Index						-0.046***
						(0.007)
Constant	1.100***	0.318*	0.412**	-0.719***	0.789***	0.653***
	(0.209)	(0.172)	(0.188)	(0.274)	(0.128)	(0.053)
Observations	41,530	41,530	41,530	41,530	41,530	41,530
R-squared	0.001	0.001	0.001	0.002	0.001	0.001
Predictors	80	80	80	80	80	80
AC + (P x AC)	-1.516	1.200	-1.297	2.909	-0.213	-0.082
p-value	0.014	0.096	0.006	0.002	0.648	0.001

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. Trading Volume 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. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Variables	Variance Rank	Trading Volume Rank	Dollar Volume Rank	Difference in Short Interest
Post-Sample (S)	0.006***	0.006***	0.010***	0.125***
	(0.001)	(0.001)	(0.001)	(0.017)
Post-Publication (P)	0.004***	0.010***	0.013***	0.354***
	(0.001)	(0.001)	(0.001)	(0.014)
Constant	0.527***	0.515***	0.445***	0.173***
	(0.001)	(0.001)	(0.001)	(0.000)
Observations	41,530	41,530	41,530	31,646
R-squared	0.407	0.434	0.520	0.450
Predictor FE?	Yes	Yes	Yes	Yes
Null: S=P	0.006	0.001	0.001	0.000

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. In-Sample Index Return is the equal-weighted return of all other unpublished predictor portfolios. Post-Publication Index Return is an equal-weighted return of all other published predictor portfolios. Standard errors are computed under the assumption of contemporaneous cross-sectional correlation between panel portfolio residuals. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Variables	Return
In-Sample Index Return	0.758***
	(0.034)
Post Publication Index Return	-0.001
	(0.010)
P x In-Sample Index Return	-0.706***
	(0.055)
P x Post-Pub. Index Return	0.626***
	(0.057)
Post-Publication (P)	-0.013
	(0.077)
Constant	0.085**
	(0.041)
Observations	35,188
Predictors	80
R-squared	0.039