Does Academic Research Destroy Stock Return Predictability?*

R. David McLean
University of Alberta and MIT Sloan School of Management
Phone: 774-270-2300
Email: rdmclean@MIT.edu

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

October 23, 2012

Abstract

We study the out-of-sample and post-publication return-predictability of 82 characteristics that are identified in the academic literature. The average out-of-sample decay due to statistical bias is about 10%, but not statistically different from zero. The average post-publication decay, which we attribute to both statistical bias and price pressure from aware investors, is about 35%, and statistically different from both 0% and 100%. Consistent with informed trading, after publication, stocks in anomaly portfolios experience higher volume, variance, and short interest, and higher correlations with portfolios that are based on published anomalies. Consistent with costly (limited) arbitrage, the post-publication return decline is greater for anomaly portfolios that consist of stocks that are large, liquid, have high dividend yields, and have low idiosyncratic risk.

* 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, Brandeis University, Boston College, CKGSB, HBS, MIT, University of Toronto, 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 Turan Bali, Shane Corwin, Mark Bradshaw, Alex Edmans, Lian Fen, Bruce Grundy, Pierluigi Balduzzi, David Chapman, Francesco Franzoni, Xiaohui Gao, Clifford Holderness, Darren Kisgen, Owen Lamont, Jay Ritter, Robin Greenwood, Paul Schultz, Bruno Skolnik, Jeremy Stein, Matti Suominen, Allan Timmermann, Michela Verado, and Artie Woodgate 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 that they provide insight into the future. Whether or not the typical relation continues outside of the study's original sample is an open question.[1] Many papers note whether a few, specific cross-sectional relations continue, although they often have contradictory messages about out-of-sample predictability.

Inference from a study of a small number of anomalies can be clouded by uncertainty of how the anomalies where selected. The publication process may favor follow-up studies that support the original study, thus the results of a particular study will not provide reliable evidence of how a typical anomaly fares post-publication. Some of the relations between characteristics and returns that we consider are consistent with rational asset pricing models, while some are anomalous. We abstain from differentiating between characteristics and refer, loosely, to all characteristics as "anomalies." This paper provides a novel attempt to synthesize information from 82 characteristics that have been shown to explain cross-sectional stock return variation in peer reviewed finance, accounting, and economics journals.

Publication may have no impact on post-publication anomalies. Cochrane (1999) contends that predictability is the outcome of rational asset pricing. Publication does not affect the information used by rational agents, so return predictability should be unaffected by publication. Examples of rational asset pricing include returns reflecting compensation for covariance with the market portfolio, such as Sharpe (1964), or compensation for transaction costs, such as Amihud and Mendelson (1986). These models assume that agents have rational expectations, and suggest that publication of the cross-sectional result should not affect the perceived marginal cost

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

that drives expected return differences. If this is the case, in- and out-of-sample return-predictability should be similar.[2]

If anomalies are purely the result of statistical biases, then we would expect them to disappear out of sample. We use the term "statistical biases" to describe a broad array of biases that are inherent to published research. Leamer (1978) investigates the impact of "specification search" biases. These biases occur if the choice of model is influenced by the data used to test the model. Lo and MacKinlay (1990) examine a specific type of specification search bias found in finance, which they refer to as the "data snooping bias." This effect arises when researchers form portfolios to show that a characteristic is related to future returns, and the choice of how the portfolios are formed is affected by the result. A second type of bias is sample selection bias, studied in Heckman (1979), where the data used for estimation is influenced by the result of the test. A third type of bias, considered by Hedges (1992), is that the publication of a study could be correlated with the magnitude of the test statistic that the study produces. A fourth type of bias occurs if a strategy's spuriously high returns attract academic attention to the strategy, and thus, making the publication date endogenous.[3] Fama (1991) also describes a publication bias, 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." To the extent that anomalies are caused by such biases, we should observe a decline in return-predictability out of sample.

Alternatively, if anomalies are the result of mispricing, and if publication draws the attention of sophisticated investors, then we might expect anomalies to disappear post-publication because of arbitrage. Using index funds formed on the basis of size and value,

[2] This logic can be extended to irrational return predictability as well. Industry research may lead academic research, such that the information in academic publications is redundant to market participants.
[3]We thank Allan Timmermann for pointing out his possibility.

Schwert (2003) shows that the size and book-to-market effect have disappeared, and interprets the findings by stating, "the activities of practitioners who implement strategies to take advantage of anomalous behavior can cause the anomalies to disappear."[4]

Others argue that practitioners will not trade enough to fully eradicate anomalies. Rather, mispricing may continue at a reduced level. Delong, Shleifer, Summers, and Waldman (1990) show than systematic noise trader risk may allow mispricing to continue. Pontiff (1996, 2006) shows that other arbitrage costs, and in particular holding costs associated with idiosyncratic risk, can prevent sophisticated trades from entirely eliminating mispricing.[5] Shleifer and Vishny (1997) point out that these effects are greater if arbitrageurs are agents who are evaluated by uninformed principals that confuse volatility with performance.

In the short run, publication may induce more acute return predictability. If characteristics are persistent over time, publication will trigger new flows of capital to characteristic portfolios, which, in turn, generate higher returns. These higher returns are temporary, since subsequent returns will be lower due to the reduced mispricing.

We conduct our analysis using 82 different anomalies from 68 different studies. The period during which an anomaly is outside of its original sample but still pre-publication, is useful for estimating the effects of statistical biases. Using this framework, we estimate the effect of statistical bias to be about 10%. Thus, an in-sample finding that implies a 5% alpha is expected to produce a bias-free alpha of at least 4.5%. This finding is statistically insignificant—we cannot reject the hypothesis that there is no statistical bias.

[4] To 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] For further evidence of this effect see Duan, Hu, and McLean (2009 and 2010) and McLean (2010)

We estimate the average anomaly's post-publication return decays by about 35%. Thus, an in-sample alpha of 5% is expected to decay to 3.25% post-publication. We attribute this effect due to both statistical biases, and to the activity of sophisticated traders who observe the finding. We can reject the hypothesis that post-publication return-predictability does not change, and we can also reject the hypothesis that there is no post-publication alpha. Combining this finding with an estimated statistical bias of 10% implies a lower bound on the publication effect of about 25%. These findings are robust to replacing publication date with Social Science Research Network (SSRN) posting date, and they do not appear to be caused by time trends in anomaly returns.

We further investigate the effects of publication by studying traits that reflect whether publication increases trading in stocks that make up an anomaly portfolio. We find that variance, turnover, dollar volume, and short interest all increase out-of-sample and post-publication. This is consistent with the idea that academic research draws attention to anomalies, which in turn increases trading in the anomaly.

As we mention above, Pontiff (1996 and 2006), and Shleifer and Vishny (1997) point out that costs and risks associated with arbitrage could prevent mispricing from being completely eliminated. If anomalies are the result of mispricing and if publication draws the attention of arbitrageurs, then the anomalies that decline the most post-publication should be those that are least costly to arbitrage. We find evidence of this effect, as the post-publication decline is greatest for anomalies that are concentrated in large market capitalization stocks, high dollar volume stocks, low idiosyncratic risk stocks, and stocks that pay dividends.

We next investigate whether academic publication is associated with changes in covariance between anomalies. We find that after a specific anomaly is published, the correlation of the anomaly portfolio with other yet-to-be-published anomaly portfolios decreases, while the correlation of the anomaly with other already-published anomaly portfolios increases. One interpretation of this finding is that anomalies are the result of mispricing, and mispricing has a common source; this is why in-sample anomaly portfolios are highly correlated. Publication causes more arbitrageurs to trade on the anomaly, which causes anomaly portfolios to become more correlated with already-published anomaly portfolios that are also being pursued by arbitrageurs, and less correlated with yet-to-be-published anomaly portfolios.

Previous studies that investigate the robustness return-predictability study a limited number of anomalies and produce conflicting messages. A few studies explicitly compare pre- and post-publication performance. For example, Schwert (2003) considers size and book-to-market index funds that were created in response to academic publications, and is unable to reject the null hypothesis of a zero alpha post-publication. Green, Hand, and Soliman (2011) find similar effects in accrual based portfolio returns after the original accrual paper was published. In contrast, Jegadeesh and Titman (2001) re-estimate momentum returns in the period after their original sample ended and show that the average return of a long-short momentum portfolio increased.[6]

Some papers are more aggressive in the number of anomalies that they analyze, but their comparisons are not based on the publication dates or sample-end dates of earlier research. Again, these papers also produce conflicting results. Haugen and Baker (1996) compare the predictive ability of 11 anomalies in two subsamples, while Chordia, Subrahmanyan, and Tong (2011) compare 7 anomalies in two more recent subsamples. Haugen and Baker show that each

[6] Another related paper is Lewellen (2011), which uses 15 variables to produce a singular cross-sectional return proxy. Lewellen's study focuses on whether estimation from 10 years rolling regressions is useful in predicting next period's cross section of returns. His study is concerned with the accuracy of estimation and not with discrete changes in predictability from publication and sample-selection effects.

of their 11 anomalies produces statistically significant returns in the second-subsample, whereas Chordia, Subrahmanyan, and Tong show that none of their 7 anomalies is statistically significant in their second-subsample.

A contemporaneous study by Green, Hand, Zhang (2012) substantially increases the field by identifying 300 published and unpublished anomalies. They argue that anomaly returns are persistent and tend to be orthogonal to other anomalies (as we confirm). Unlike our paper, their paper does not estimate anomaly decay parameters as a function of publication or sample-end dates.

1. Research Method

We identify studies that suggest cross-sectional relations between variables that are known in a given month and stock returns in the following month(s). We limit ourselves to studies in the academic peer-reviewed finance and accounting literatures 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 for which the long and short side of the portfolio are determined by the percentile placement of each stock relative to the universe of stocks that are included in both Compustat and CRSP. Despite the fact that the Fama-

MacBeth slopes are regression slope coefficients, they are also long-short portfolio returns with positive exposure to the underlying characteristic (Fama, 1976).

Some of the studies that we identify demonstrate the univariate relation between 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 82 cross-sectional relations from 68 different studies. The study with the most number of original cross-sectional relations (4) that we utilize is Haugen and Baker's 1996 study of cross-section stock returns in the *Journal of Financial Economics*. Haugen and Baker (1996) investigate more than four cross-sectional relations, but some of these relations were documented by other authors earlier, and are therefore associated with other publications in our study. The first study in our sample is Blume and Husic's 1972 *Journal of Finance* study of how price level relates to future stock returns. The most recent study is Bali, Cakic, and Whitelaw's (2011) *Journal of Financial Economics* study that shows that the maximum daily return that a security experiences in the preceding month predicts the next period's monthly return.

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*. These studies argue that expected cross-sectional returns are a rational response to risk and costs, and thus sophisticated traders that observe the findings of these papers should not be induced to trade.

For each study, we keep track of the official date that the journal was released, and when available, the date that the paper was first released on SSRN. We also keep track of the beginning and ending month of the sample period that was used to produce the original findings of cross-sectional predictability.

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, but it does cover S&P ratings, and so we use S&P rating downgrades instead.

Each characteristic that uses accounting data is winsorized, such that values that are below the 1st percentile are assigned the value of the 1st percentile, and values that are above the 99th percentile are assigned the value of the 99th percentile. We compute two predictability statistics. First, we calculate 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. Second, we calculate the return of a portfolio that each month invests in stocks in the top 20th percentile of the characteristic minus the return of a portfolio that invests in stocks in the bottom 20th percentile of the characteristic.

2. Creating the Data and In-Sample Replicability

Summary statistics for the anomalies that we study are provided in Table 1. For most of our results, we define 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 a robustness check, we also consider publication date to be the earlier of the actual publication date and the first time that paper appeared on the SSRN. For this measure, the average number of months between the end of the sample and publication date is 44 months. For comparison, the average original in-sample span is 323 months.[7] Our average out-of-sample span is 139 months.

For all anomalies, we calculate monthly Fama-MacBeth slope coefficients. Fifteen of our 82 characteristics involve binary variables, such as dividend initiation (Michaely, Thaler, and Womack, 1995). For all anomalies that do not involve a binary characteristic, we also calculate the long-short portfolio monthly return from the extreme quintiles. These portfolios are equally weighted unless the primary study presents value-weighted portfolio results. For example, value weighting is used for the idiosyncratic risk results that follow (Ang, Hodrick, Xing, and Zhang, 2006).

Our goal is not to perfectly replicate a paper. This is impossible since CRSP data changes over time and often papers omit details about precise calculations. In some cases, our failure to calculate suitable t-statistics is likely caused by our aggressive efforts to include characteristics. For example, we are unable to calculate significant t-statistics for the IPO effect (Ritter, 1991) and the dividend yield measure of Naranjo, Nimalendran, and Ryngaert (1998). For the case of IPOs, as previously mentioned, we use the initiation of CRSP coverage as a proxy for IPOs. For

[7] A discussion of characteristic computations, and a list of articles, publication dates, and sample dates, should be available in an appendix in the next version of the paper.

the case of dividend yield, the original paper used additional data that is not included in CRSP and Compustat.

Ten of our average in-sample Fama-MacBeth slope coefficients produce t-statistics that are between -1.25 and 1.25.[8] We do not include these characteristics in the paper's main tests. Thus, a total of 72 (82 – 10) characteristics are used in the paper's primary tests. Admittedly, the decision to use a t-statistic cut-off of 1.25 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 1.25 seemed reasonable to us. That stated, only three of the 72 characteristics that we include in the paper's analyses have t-statistics that are less than 1.80.

2.1. Preliminary Findings

Table 2 reports anomaly-level summary statistics regarding the out-of-sample and post publication return-predictability of the 72 characteristics that we were able to replicate. To be included in the tests in this table, we require that an anomaly have at least 36 observations during the measurement period (e.g., post-publication). We relax this restriction in our panel regression tests which weight each anomaly-month observation equally, rather than each anomaly.

To estimate the statistics in Table 2, we first calculate the in-sample mean for each anomaly, as described in the previous section. We then scale the out-of-sample and post-publication coefficients by the in-sample means. We use the scaled values to generate statistics that reflect the average out-of-sample and post-publication return of each anomaly relative to its

[8] If a characteristic is not associated with a t-statistic outside of the -1.25 to 1.25 range, both co-authors independently wrote code to estimate the effect.

in-sample mean. We generate individual statistics for each anomaly and then take a simple average across all of the anomalies. We do this for the continuous version of each anomaly (Panel A), the portfolio version (Panel B), and the strongest form (continuous vs. portfolio) (Panel C).

As an example, Panel A shows that if we use the continuous estimation of each anomaly (continuous variable in the Fama and MacBeth (1973) regression), then the average anomaly's return is 78% of its in-sample mean during the out-of-sample, but pre-publication period. However, this decline is not statistically significant (*t*-statistic = -1.40). Once published, the average anomaly's return is only 51% of its in-sample mean and this decline is highly significant (*t*-statistic = -4.91). The results are similar throughout the 3 panels, which is the continuous, portfolio, and strongest form versions of the anomalies.

Because these results are summarized at the anomaly level, the statistics give more weight to observations from anomalies that have shorter sample periods. As an example, the size effect (Banz, 1981) has monthly observations that go back to 1926, while the distress effect (Dichev, 1998), which uses credit ratings data, begins in 1981. Hence, if we equal-weight each anomaly, as we do in Table 2, then one observation from the distress anomaly gets a much larger weight than does one observation from the size anomaly. Also, the statistics in Table 2 do not consider correlations across anomaly portfolios. The random effects regressions that we describe below produce results that are robust to these issues, although both sets of tests find an insignificant decline out-of-sample and a significant decline post-publication.

3. Main Results

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

We now more formally study the return-predictability of each characteristic relative to its original sample period and publication date. We first introduce our regression methodology, which utilizes random effects, to control for cross-anomaly portfolio correlations. In the discussion that follows, PR_{it} denotes the portfolio return associated with characteristic i in month *t*; this is either the Fama-MacBeth slope coefficient from a regression of monthly returns on characteristic i in month t or the return from a long-short portfolio in the characteristic's extreme quintiles formed by monthly sorts on the characteristic.

We first compute the average portfolio return for each anomaly i, using the same sample period as in the original study. This average will be expressed as $\overline{PR_i}$. The next step is to normalize each monthly portfolio return by scaling the observation by $\overline{PR_i}$. This normalized portfolio return will be denoted $\widetilde{PR}_{it}$. In order to document changes in return predictability from in-sample to out-of-sample, and from in-sample to the period that is both out-of-sample and post-publication, we estimate the following equation:

$$\widetilde{PR}_{it} = H_{int} + H_{post-sample}\, D_{it}^{post-sample} + H_{post-pub} D_{it}^{post-pub} + e_{it} \qquad (1)$$

In this equation, $D_{it}^{post-sample}$ is a dummy variable that is equal to one if month *t* is after the end of the original sample but still pre-publication, and zero otherwise, while and $D_{it}^{post-pub}$ is equal to 1 if the month is post-publication, and zero otherwise. e_{it} is the residual from the estimation and the *H* variables are slope coefficients.

For the basic specification in equation (1), the intercept, H_{int}, will be very close to unity. This occurs since the average normalized portfolio return that is neither post-sample nor post-publication is unity by construction—the normalized return is the actual in-sample return dividend by the in-sample average. This accomplishes the objective of allowing us to interpret the slopes on the dummy variables as percentage decays of in-sample slopes. A benefit of using returns that are normalized by the in-sample mean as opposed to using the in-sample mean as an independent variable, is that this enables us to use a longer time-series to estimate the appropriate variance-covariance matrix. As we mention above, the portfolio returns from the same month are likely to be correlated and, because of this, we estimate equation (1) with random-effects by characteristic portfolios. In addition, we also cluster our standard errors on time.

The coefficient, $H_{post-sample}$, estimates the total impact of statistical biases on anomaly performance (under the assumption that sophisticated traders are unaware of the working paper before publication). $H_{post-pub}$ estimates both the impact of statistical biases and the impact of publication. If statistical biases are the cause of in-sample predictability, then both $H_{post-sample}$ and $H_{post-pub}$ should be equal to -1. Such a finding would be consistent with Fama's (1991) conjecture that return anomalies in academic studies could be the outcome of data-mining. If anomalies are the result of mispricing, and arbitrage resulting from publication corrects all mispricing, then $H_{post-pub}$will be equal to -1, and $H_{post-sample}$ 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 $H_{post-sample}$ and $H_{post-pub}$ should equal zero. This last case would also be consistent with Cochrane's (1999) observation that if return-predictability documented in academic papers reflects risk, then it should not change outside of the original sample period.

3.2. Anomaly Dynamics Relative to End of Sample and Publication Dates

Table 3 presents regression estimates of how predictability varies through the life-cycle of a publication. The first column, labeled "Continuous," uses Fama-MacBeth coefficients (again, scaled by in-sample means) generated from regressions that use continuous measures of the characteristic (e.g., size or past returns) as the dependent variable. The results suggest that between the end of the sample and the publication date, the magnitude of the long-short returns fall, on average, by about 20%. We are unable to reject the hypothesis that this drop is statistically significant from zero. Post-publication, the decline is 42% and statistically significant from both 0 and -100%. Thus, cross-sectional predictability continues post-publication at a significant, albeit muted level.

The third column uses either the Fama-MacBeth coefficients from regressions that use continuous variables, or long-short extreme quintile portfolios, depending on which method produces the highest in-sample statistical significance. In this regression the post-publication decline is estimated to be 37%, which is similar to the slopes (42% and 35%) estimated in the regressions that use the continuous and portfolio estimates respectively. If the original cross-sectional relations were purely noise, then selecting a weighting method based on in-sample significance would produce the largest decay in post-sample returns, however this is not the case.

The fourth column considers an alternative publication date that is based on either the actual publication date or the first SSRN posting date, whichever is earlier. In this regression, the post-publication coefficient estimates a decay of 34%, showing that small changes to publication dates do not have an effect on the findings. This finding makes sense, since the post-publication coefficient is essentially a test of the difference between the in-sample and post-publication values of the normalized portfolio returns. We have a total of 9,984 post-publication portfolio-month returns, which increases to 10,797 if we instead use the SSRN posting date as the publication date. Hence, this change in definition increases the post-publication sample by only 7.5%, which is not a large difference.

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 four regressions, the post-publication coefficient is significantly larger at the 10% levels. Hence, the decline in return-predictability that is observed post-publication exceeds 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.

3.3. A Closer Look at Anomaly Dynamics

Table 4 considers a random effects regression that expands on Eq. (1). The regression reported in Table 4 consists of 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.

In Table 4, the coefficient for the last 12 months of the sample period is negative and insignificant. Similarly, the coefficient for the first 12 months out-of-sample is positive and insignificant. If researchers choose where to end their samples with the purpose of getting stronger results, then we expect the exact opposite signs on both coefficient. Hence, these results run counter to claims that authors opportunistically select sample end dates.

The out of sample but pre-publication coefficient and the coefficients for the first 2 years out of sample are all negative and similar in magnitude (-0.178 to -0.292). One interpretation of these effects is that some arbitrageurs learn about these anomalies before publication, however even in the first couple of years after publication this information is not widely used.

The coefficients for post-publication years 3, 4, and 5 demonstrate the biggest decay in predictability. For this time period, predictability is about half of what it is in-sample. The coefficient for all months after the fifth year is -0.307. Hence, anomalies appear to make large declines during years 3-5, and then partially recover thereafter, albeit still at a lower level than in-sample.

3.4. Publication Effect or Time Trend?

Table 5 considers alternative explanations for the results in Tables 2 and 3. 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 (Pontiff, 1996). 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. Chorida, Subrahmanyam, and Tong (2011) show that in the 1993 to 1999 time period, ten anomalies that were previously associated with cross-sectional returns failed to achieve statistical significance. They attribute this result to lower transaction costs and more trading activity from informed traders. Hence, it could be the case that anomalies are diminishing because the costs of trading on these anomalies have declined over time.

We consider these alternative explanations with three time series variables. First, we construct a time variable that is equal to 1 in January 1926 and increases by 1 during each consecutive month in our sample. If transactions costs have been decreasing linearly over time, this variable should be negatively associated with anomaly returns. Our second variable is an anomaly-specific time variable that is equal to 1 during the first month after publication and increases by 1 in each subsequent month. If sophisticated traders learn about anomalies slowly and linearly after publication, then this should share a negative relation with anomaly returns. The coefficients and standard errors for both of these monthly-time variables are reported in percent. Lastly, we use a post-1993 indicator variable to proxy for the discrete bifurcation of the data that Chordia et al. use.

The regressions in Table 5 use the returns for the method (continuous vs. portfolio) that produces the most statistically significant in-sample anomaly returns. In regressions where the time variables are used without other regressors (columns 1-3), both months after 1926 (time) and months after publication have negative slope coefficients that are significant at the 1% level. The post 1993 indicator variable is -12.2%, which is consistent with Chordia et al, however the p-value is 0.255, exceeding the typical bound of significance.

Columns 4-6 include both the time variables, and the post-sample and post-publication indicators. As in Table 3, the slope of the post-sample indicator is negative and insignificant for all specifications, while the slope on the post-publication indicator is negative and significant in all of the regressions. The magnitude of the post-publication slope ranges from about -29% to -43%, similar to what is reported in Table 3. Thus, this evidence implies that post-publication changes in predictability dominate trends in predictability, which is insignificant in the presence of the post-publication indicator.

3.5. Publication and Trading in Anomaly Stocks

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, such as turnover, dollar value of trading volume, stock return variance, and short interest. 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. Short interest is shares shorted scaled by shares outstanding.

To estimate these effects we perform monthly ranks of all of the stocks in CRSP based on these four different measures. We focus on rankings because these trading activity measures are likely to have market-wide time-trends (e.g, turnover is, on average, higher now as compared to 1930). For each anomaly-month, we compute the average ranking among the stocks that enter either the long or the short side of the anomaly portfolio. We scale each anomaly-month's ranking by the anomaly's in-sample average and test whether the ranking changes out-of-sample and post-publication.

We report the results from these tests in Tables 6 and 7. Table 6 is similar to Table 2, in that it summarizes the effects at the anomaly level. Like Table 2, we require that an anomaly have at least 36 observations during the measurement period (e.g., post-publication). As before, we relax this restriction in our regression tests reported in Table 7, which weight each anomaly-month observation equally. Table 6 shows that turnover, dollar volume, variance, and short interest share similar reactions. Once the sample period of the academic paper ends, all of the trading activity measures have statistically significant increases. The effects with the trading characteristics are significant in both the post-sample but pre-publication period, and in the post-publication period. The total change between the original sample period and the post-publication period ranges from 9% to 22% for the different trading measures. Academic research seems to have some effect on trading activity, and it occurs before the paper is published. Our results regarding short interest mimic the results in a contemporaneous paper by Hwang and Liu (2012), who use short interest as a proxy for the popularity of 11 anomalies with sophisticated traders. They compare short interest around academic publication and find similar results.

Similar to Table 3, Table 7 estimates a regression akin to Eq. (1); only the dependent variable is the normalized rank of the trading characteristic, rather than the normalized return. Lagged values of the dependent variables are included in the regressions since the trading activity measures are persistent through time. The coefficients are reported in percent.

Similar to the results reported in Table 6, the results in Table 7 show statistically significant slopes on the post-sample dummy variables for all specifications. The interpretation of these effects is somewhat different from Table 2. In Table 2, return evidence of informed trading is not apparent until the publication date. The Table 7 findings imply informed trading occurs more quickly. This could be caused by an increased recognition of the anomaly resulting from circulation of the paper and conferences. A further increase in trading-based activity is observed post-publication, in that the coefficient is positive and significant in all of the regressions. For variance and short interest, the slope on post publication is higher and significantly different from the slope on post-sample. As in Table 3, there appears to be an extra effect from publication with these variables. We cannot reject the null that the slopes are equal in dollar volume and turnover regressions.

3.5. Costly Arbitrage and Anomaly Persistence Post-Publication

Some of the results in the previous tables are consistent with the idea that publication attracts arbitrageurs, which results in smaller anomaly 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, anomaly portfolios that consist 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 anomalous returns are the outcome of rational asset pricing, we would not expect the post-publication decline to be related to arbitrage costs. Keep in mind that our returns are scaled by the in-sample mean, so a decline implies that the returns shrink towards zero. Anomalies that produce negative returns have an increase in returns, while anomalies that produce positive returns experience a decrease in returns.

Previous papers in the limited arbitrage literature estimate differences across stocks within an anomaly portfolio (see Pontiff, 2006; Duan, Hu, and McLean, 2010; and McLean, 2010). In contrast, we estimate differences across anomaly portfolios, and there is less variation in traits such as size and variance across anomaly portfolios than across individual stocks within an anomaly portfolio. Hence, our cross-anomaly costly arbitrage tests might be less powerful than previous costly arbitrage tests. Another difference between our test and previous literature is that previous studies assume rational expectations of the informed traders through-out the entire sample. In this framework, the informed trader had knowledge of the anomaly before (and after) the publication date. Our current test assumes that publication date provides information to some sophisticated traders, which, in turn, causes variation in anomaly decay post-publication.

In Table 8, the dependent variable is the normalized anomaly return, limited post-publication months. To create the independent variables, we perform monthly ranks of all of the stocks in CRSP based on 3 transaction cost measures; size, dollar volume, and bid-ask spreads, and two holding costs measures: idiosyncratic risk and a dividend-payer dummy.

Idiosyncratic risk is a holding cost since idiosyncratic risk is incurred every period the position is open. 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 residual and the previous day's residual. The monthly measure is created by adding up the daily data from a given month. Dividends mitigate holding costs since they decrease the effective duration of the position.[9] We use a dummy variable equal to unity if a firm paid a dividend and zero otherwise. 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). Stocks with high dollar volume and low spreads are more liquid, and should therefore be easier to arbitrage.

For each anomaly-month, we compute the average ranking among the stocks that are in either the long or the short side of the anomaly portfolio. We create an anomaly-month average for each trait and then take an average of the monthly averages to come up with a single in-sample-anomaly average. We measure the traits in-sample, as it could be the case that trading caused by publication has an effect on the variables.

[9] For more detail, Pontiff (2006) derives the impact of these effects on sophisticated investors.

Table 8 shows that there are significantly larger post-publication declines for anomaly portfolios that are subject to lower arbitrage costs. All coefficients have the expected sign and are statistically significant. The holding cost variables command the lowest p-values. Consistent with holding costs affecting sophisticated traders, high idiosyncratic risk portfolios decay less, while portfolios with dividend paying stocks decay more.

3.6. The Effects of Publication on Correlations between Anomaly Portfolios

In this section, we study the effects that publication has on the correlation between anomaly portfolios. If anomalies reflect mispricing, and if mispricing has common causes, then we might expect in-sample anomaly portfolios to be correlated with other in-sample anomaly portfolios. This might also occur if some investors specialize in trading in anomalies that have not been uncovered by academics. If publication causes arbitrageurs to trade in an anomaly, then it could cause an anomaly portfolio to become more highly correlated with other published anomalies and less correlated with unpublished anomalies.

In Table 9, each anomaly-month return is regressed on an equal-weighted portfolio of other anomalies that are pre-publication and an equal-weighted portfolio of other anomalies that are post-publication. We include a dummy variable that indicates whether the anomaly is post-publication, and interactions between this dummy variable and the pre-publication and post-publication anomaly portfolios returns. As before, the monthly anomaly returns are scaled by their in-sample mean.

We report results for the continuous (Panel A), portfolio (Panel B), and strongest form (Panel C) versions of each anomaly. The results for the three specifications are similar, so we focus our discussion on the strongest form results in Panel C. The results show that when an anomaly is pre-publication, the associated anomaly portfolio returns are significantly related to the returns of other pre-publication anomaly portfolios. The slope coefficient is 0.634 and its p-value is 0.00. In contrast, the slope coefficient or beta of a pre-publication anomaly returns with anomaly returns that are post-publication is 0.025.

The interactions show that once an anomaly is published, the returns on portfolios that focus on the anomaly are less correlated with the returns of other pre-publication anomaly portfolios, and more correlated with the returns of other post-publication anomaly portfolios. The slope on the interaction of the post-publication dummy with the return of the portfolio consisting of in-sample anomalies is -0.555 (p-value = 0.00). Hence, once an anomaly is published, the correlation of its respective portfolio returns with the returns of other yet-to-be-published anomaly returns virtually disappears, as the overall coefficient reduces to 0.634 – 0.555 = 0.079. The slope on the interaction of the post-publication dummy and the returns of the other post-publication anomalies is 0.399 (p-value = 0.00), so there is a significant correlation between the portfolio returns of a published anomaly and other published anomalies. Consistent with Green et al. (2012), the R^2 between anomalies ranges from 6% to 20%, thus there are benefits for sophisticated traders to consider multiple anomalies.

3.6. Predicting Predictability

In this section of the paper, we ask whether anomaly portfolios exhibit persistence. 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 suggest that momentum, or perhaps shorter-term persistence, might exist among our larger sample of anomalies.

We estimate a random-effects regression of the normalized anomaly return (based on the strongest form) on a constant and three lags of normalized returns. We estimate this model separately in- and out-of-sample. We report these results in Table 10. Except for the expected difference in intercepts, both autoregressions are similar. Hence, publication does not seem related to the short-term return autocorrelation. There is strong positive, and significant, autocorrelation at the monthly level. The second and third autocorrelations are close to zero and insignificant. These results imply that last month's long-short returns help predict the current month's return, although there does not seem to be persistence that continues past one month.

The last 2 columns in Table 10 consider two types of momentum, the sum of last 6 months' long-short returns and the sum of the last 12 months' long-short returns. If our portfolio returns were on long-only or short-only portfolios, we would have made the decision to use compound returns. Long-short returns may create unlimited liability, making compounding problematic. We think that summing long-short returns is a reasonable way to avoid this problem, while still capturing momentum. The slope coefficients on both 6-month and 12-month momentum are positive, and the 12-month coefficient is significant. Taken in their entirety, the results in Table 10 show that anomaly returns are persistent, which is broadly consistent with the findings of Moskowitz et al.

4. Conclusions

We investigate the out-of-sample and post-publication return predictability of 82 characteristics that have been shown to predict cross-sectional returns by academic publications in peer-review journals. We estimate an upper bound estimate of statistical bias of about 10%, although we are unable to reject the null of no bias. We estimate post-publication decay to be about 35% and we can reject the hypothesis that there is no decay. We can also reject the hypothesis that the cross-sectional predictive ability disappears entirely. This finding is better explained by a discrete change in predictive ability, rather than a declining time-trend in predictive ability.

In addition to the finding that anomaly portfolio returns decline post-publication, several of our other findings are consistent with the idea that academic research draws attention to anomalies, which results in more trading in anomaly stocks. First, trading activity in these stocks changes out-of-sample and post-publication, as variance, turnover, dollar volume, and short interest all increase significantly out-of-sample and post-publication. Second, anomaly portfolios that are subject to higher transactions and holding costs have lower declines post-publication. This is consistent with the idea that arbitrage costs limit arbitrage and protect mispricing. Finally, we find that before a characteristic is featured in an academic publication, the returns of the corresponding anomaly portfolio are highly correlated to the returns of other portfolios of yet-to-be-published anomaly stocks. After publication, the sensitivity to yet-to-be-published anomaly portfolios returns decreases and the sensitivity to already-published anomaly portfolios returns increases.

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.

Bali, Turan G., Cakici, Nusret, and Whitelaw, Robert F., 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.

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

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.

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., 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, Francsco 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.

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..

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

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

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

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.

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.

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

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.

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

Table 1. Summarizing the characteristics in and out-of-sample.

This table reports summary statistics for the 82 different return-predicting characteristics studied in this paper. The second column reports the number of characteristics that fit the criteria described in the first column, and that number as a percentage of the total number of characteristics in parentheses. The return-predictability of each characteristic is estimated using monthly Fama and MacBeth (1973) regressions. Each continuous characteristic is estimated twice; once using a continuous variable, and once using a portfolio variable that is equal to 1 if the stock is in the buy quintile, -1 if the stock is in the sell quintile, and zero otherwise.

Total number of return-predicting characteristics:	82
Characteristics from Finance journals	61 (74%)
Characteristics from Accounting journals	19 (24%)
Characteristics from Economics journals	2 (2%)
Characteristics that are binary (e.g. credit rating downgrade):	15 (18%)
Characteristics that are continuous (e.g. size):	67 (82%)
Characteristics that we could replicate in-sample:	72 (88%)
Replicated, continuous characteristics that are stronger as a continuous variable	36 (50%)
Replicated, continuous characteristics that are stronger as a quintile portfolio variable	36 (50%)

Table 2. Summarizing the out-of-sample and post-publication return predictability of the characteristics.

This table reports summary statistics for the out-of-sample and post-publication return predictability of the 82 replicated return-predicting characteristics used in this paper. To be included in these tests the characteristic had to both be replicated in-sample, and have at least 36 observations in the out-of-sample or post-publication measurement period. Each characteristic is estimated using a monthly Fama and MacBeth (1973) regression. Each continuous characteristic is estimated twice, first using a continuous variable, and then using a portfolio variable that is equal to 1 if the stock is in the long quintile, -1 if the stock is in the sell quintile, and zero otherwise. We estimate the in-sample mean coefficient for each characteristic, and then scale each monthly coefficient by the in-sample mean. We then take averages of the scaled coefficients during the out-of-sample and post-publication periods for each anomaly, average the averages across anomalies, and report them in the Table below. A value of 1 means the average coefficient is the same during the in-sample and out-of-sample period. A value of less than 1 (greater than 1) means the return-predictability declined (increased) out-of-sample. The t-statistic tests whether the reported value is equal to 1.

Panel A: Continuous Regressors	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	0.78	0.51
Standard Deviation	1.22	0.81
t-statistic	-1.40	-4.91
Percentage <1	63%	82%
Anomalies Included	60	66

Panel B: Portfolio Regressors	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	0.90	0.47
Standard Deviation	1.29	1.20
t-statistic	-0.58	-3.62
Percentage <1	57%	68%
Anomalies Included	60	66

Panel C: Strongest Form	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	0.77	0.51
Standard Deviation	1.16	0.97
t-statistic	-1.56	-4.02
Percentage <1	65%	78%
Anomalies Included	60	66

Table 3. Regression of long-short characteristic based returns on time indicator variables.

This regression models the return-predictability of each characteristic over time, relative to its original sample period and publication date. The monthly return for each characteristic is generated via a monthly Fama and MacBeth (1973) regression. Each monthly coefficient is scaled by the characteristic's mean coefficient during the study's original sample period. This scaled variable is the dependent variable, and it is regressed on dummy variables that signal whether the month is out of sample but pre-publication, and post-publication. *Post Sample* equals 1 if the month is after the end of the sample, but pre-publication. *Post Publication* is equal to 1 if the month is after the official publication date. *Post SSRN* is equal to 1 if the month is either after the official publication date, or if the month is after the first month that the study is available on SSRN. All indicator variables are equal to 0 if they are not equal to 1. The regression labeled *Continuous* uses Fama-MacBeth slopes that are generated using continuous variables. The regression labeled *Portfolios* uses Fama-MacBeth slopes from long-short quintile portfolios. The regression labeled *Strongest* uses either *Continuous* or *Portfolios* returns, depending on which method produces stronger in-sample statistical significance. P-values are in brackets for the hypothesis that the coefficient equals 0. In the three bottom rows we report p-values from Chi-Squared tests of the hypotheses that the post-sample and post-publication coefficients are equal, and that each of the coefficients is equal to -1. The regressions include random effects. Standard errors are clustered on time.

	Continuous	*Portfolios*	*Strongest*	*Strongest*
Post Sample	-0.202	-0.015	-0.097	-0.102
	(0.119)	(0.124)	(0.112)	(0.119)
	[0.090]	[0.902]	[0.386]	[0.389]
Post Publication	-0.422	-0.347	-0.369	
	(0.095)	(0.112)	(0.093)	
	[0.000]	[0.002]	[0.000]	
Post SSRN				-0.343
				(0.079)
				[0.000]
Constant	0.986	1.040	0.982	0.961
	(0.071)	(0.084)	(0.070)	(0.062)
	[0.000]	[0.000]	[0.000]	[0.000]
R^2	0.000	0.000	0.000	0.000
Observations	37,676	37,676	37,676	37,676
HO: PP-PS=0	0.073	0.010	0.020	0.050
HO: PS=-1	0.000	0.000	0.000	0.000
HO: PP=-1	0.000	0.000	0.000	0.000

Table 4. A closer look at the effects of post-sample and post-publication

This regression models the return-predictability of each characteristic relative to its original sample period and publication date. The monthly return for each characteristic is generated via a monthly Fama and MacBeth (1973) regression. We use either continuous variables or quintile portfolios based on the variables to generate the coefficient, depending on which method has stronger in-sample statistical significance. Each monthly coefficient is scaled by the characteristic's mean coefficient during the study's original sample period. This scaled variable is the dependent variable, and it is regressed on dummy variables that signal the position of the month in time relative to the study's original sample period and the study's publication date. The regressions use either continuous or portfolio returns based on which return has stronger in-sample statistical significance. *Last 12* is equal to 1 if the month is during the last year of the original sample period. *First 12* is equal to 1 during the first 12 months subsequent to the end of the original sample period. *Post First 12* equals 1 if the month is after the end of the sample, and after the first 12 months subsequent to the end of the original sample period, but pre-publication. *P1-12* is equal to 1 during the first 12 months after the official date publication date. *P13-24* is equal to 1 during months 13-24 after the publication date. *P25-36* is equal to 1 during months 25-36 after the publication date. *P37-48* is equal to 1 during months 37-48 after the publication date. *P49-60* is equal to 1 during months 49-60 after the publication date. *P>60* is equal to 1 during all months after 60 months after the publication date. All indicator variables are equal to 0 if they are not equal to 1. The regressions include random effects. Standard errors are clustered on time.

	Coefficient	**Standard Error**	**P-value**
Last 12	-0.091	0.218	0.678
First 12	0.338	0.209	0.107
Post First 12	-0.292	0.119	0.014
P1-12	-0.283	0.217	0.191
P13-24	-0.178	0.226	0.430
P25-36	-0.577	0.237	0.015
P37-48	-0.563	0.241	0.020
P49-60	-0.481	0.222	0.030
P>60	-0.307	0.090	0.001
Constant	0.964	0.063	0.000
R^2	0.000		
N	37,680		

Table 5. Regression of long-short characteristic based returns on time indicator variables and continuous time variables.

This regression models the return-predictability of each characteristic over time, and relative to the characteristic's original sample period and publication date. The monthly return for each characteristic is generated via a monthly Fama and MacBeth (1973) regression. We use either continuous variables or quintile portfolios based on the variables to generate the coefficient, depending on which method has stronger in-sample statistical significance. Each monthly coefficient is scaled by the characteristic's mean coefficient during the study's original sample period. This scaled variable is the dependent variable, and it is regressed on dummy variables that signal whether the month is out of sample but pre-publication, and post-publication. *Post Sample* equals 1 if the month is after the end of the sample, but pre-publication. *Post Publication* is equal to 1 if the month is after the official date publication date. *Time* is the number of months post-Jan. 1926. *Time Post-Publication* is the number of months post-publication. The time coefficients and standard errors are reported in percent. *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. *Spreads* is the average bid-ask spread. Fama-MacBeth slopes from either continuous variables, or long-short extreme quintiles, are used based on which return has stronger in-sample statistical significance. P-values are in brackets for the hypothesis that the coefficient equals 0. In the three bottom rows, we report p-values from Chi-Squared tests of the hypotheses that the post-sample and post-publication coefficients are equal, and that each of the coefficients is equal to -1. The regressions include random effects. Standard errors are clustered on time.

Dependent Variable is Portfolio Return Scaled by its In-Sample Mean						
Post Sample				-0.029	-0.025	-0.069
				(0.119)	(0.127)	(0.128)
				[0.806]	[0.845]	[0.588]
Post Pub.				-0.365	-0.293	-0.425
				(0.094)	(0.127)	(0.098)
				[0.000]	[0.021]	[0.000]
Time	-0.041			0.003		
	(0.000)			(0.000)		
	[0.021]			[0.900]		
Time Post Pub.		-0.190			-0.070	
		(0.000)			(0.001)	
		[0.000]			[0.280]	
Post 1993			-0.122			0.099
			(0.107)			(0.122)
			[0.255]			[0.419]
Constant	1.174	0.921	0.924	0.952	0.970	0.951
	(0.136)	(0.061)	(0.072)	(0.151)	(0.066)	(0.074)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R^2	0.000	0.000	0.000	0.000	0.000	0.000
Obs.	37,680	37,680	37,680	37,680	37,680	37,680

Table 6. Out-of-sample and post-publication trading activity among stocks in the characteristic portfolios.

This table reports summary statistics regarding the out-of-sample and post-publication trading of the 82 return-predicting characteristics that we study. To be included in these tests the characteristic had to both be replicated in-sample and have at least 36 observations in the measurement period. For each stock in each long-short portfolio, we compute its percentile ranking relative to all stocks in CRSP based on share turnover, dollar volume, and variance. We also compute the percentile ranking of short interest for stocks in the short-side or the portfolio. We then estimate an average ranking for each portfolio-month. Each monthly average is then scaled by the mean of the monthly averages during the characteristic's original sample period. Hence, a value of 1 during the post-publication period means the average monthly rank is the same during the in-sample and post-publication periods. A value of less than 1 (greater than 1) means the rank declined (increased) post-publication relative to in-sample. The t-statistics test whether the value is equal to 1.

Panel A: Turnover	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	1.11	1.11
Standard Deviation	0.19	0.16
t-statistic	4.66	5.71
Percentage >1	83%	83%
Anomalies Included	60	66

Panel B: Dollar Volume	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	1.13	1.11
Standard Deviation	0.22	0.18
t-statistic	4.58	4.64
Percentage >1	80%	73%
Anomalies Included	60	66

Panel C: Variance	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	1.10	1.22
Standard Deviation	0.16	0.16
t-statistic	4.71	11.51
Percentage >1	65%	89%
Anomalies Included	60	66

Panel D: Short Interest	Out of Sample but Pre-Publication	Post Publication
Average Scaled Coefficient	1.04	1.09
Standard Deviation	0.06	0.09
t-statistic	5.42	8.01
Percentage >1	82%	92%
Anomalies Included	56	61

Table 7. Regression of relative trading differences for portfolio stocks

This regression models the dynamics of the traits of stocks in each characteristic portfolio, relative to the characteristic's original sample period and publication date. For each stock in each long-short portfolio, we compute its percentile ranking relative to all monthly stocks based on variance, share turnover, and dollar value of volume. We also compute the percentile ranking of short interest for stocks in the short-side or the portfolio. We then generate a monthly stock-average for each anomaly. Each monthly average is scaled by the mean of the monthly averages during the characteristic's original sample period. *Post Sample* is equal to 1 if the month is after the end of the sample, but pre-publication. *Post Publication* is equal to 1 if the month is after the official date of the journal that published the study. The regressions include random effects. Standard errors are clustered on time and reported in parentheses. The coefficients are reported in percent. P-values are in brackets for the hypothesis that the coefficient equals 0.

	Variance	**Turnover**	**Dollar Volume**	**Short Interest**
Post Sample	1.108	0.300	0.281	0.867
	(0.150)	(0.118)	(0.104)	(0.100)
	[0.541]	[0.011]	[0.007]	[0.000]
Post Publication	1.991	0.281	0.262	1.893
	(0.541)	(0.149)	(0.131)	(0.148)
	[0.000]	[0.059]	[0.046]	[0.000]
Lag of Dep. Variable	51.035	97.802	98.248	81.984
	(4.403)	(1.064)	(0.864)	(0.786)
	[0.000]	[0.000]	[0.000]	[0.000]
Constant	48.970	2.238	1.779	18.047
	(0.044)	(1.224)	(1.032)	(0.732)
	[0.000]	[0.067]	[0.085]	[0.000]
R^2	0.238	0.960	0.983	0.704
Obs.	38,620	38,620	38,620	27,171
Chi Sq. Test: PP-PS	0.099	0.850	0.830	0.000

Table 8: Arbitrage costs and the persistence of characteristic return predictability

This regression tests whether different stock-traits are associated with a characteristics' change in return-predictability post-publication. The sample is limited to post-publication months. The dependent variable is the monthly long-short return of a characteristic scaled by its monthly in-sample mean. This monthly return is generated via a monthly Fama and MacBeth (1973) regression. We use either continuous variables or quintile portfolios based on the variables to generate the characteristic's monthly return, depending on which method has stronger in-sample statistical significance. The independent variables reflect various traits of the stocks in each characteristic portfolio. Each characteristic portfolio contains stocks in the highest and lowest quintiles, based on a contemporaneous ranking of the characteristic (e.g., momentum or accruals). To measure the traits of the stocks within the 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 size rank. We then take the average rank of all of the stocks in the characteristic portfolio for that month. Then, for each characteristic, we take an average of its portfolio'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 size rank of the stocks in the characteristic portfolio during the in-sample period for the characteristic. The regressions include random effects. Standard errors are clustered on time, and are reported in parentheses. P-values are in brackets for the hypothesis that the coefficient equals 0.

	Size	**Spreads**	**Dollar Volume**	**Idiosyncratic Risk**	**Dividends**
Coefficient	-1.490	0.999	-1.671	4.054	-1.381
	(0.598)	(0.592)	(0.642)	(0.855)	(0.352)
	[0.013]	[0.092]	[0.009]	[0.000]	[0.000]
Constant	1.442	0.176	1.380	-1.420	1.439
	(0.339)	(0.262)	(0.296)	(0.430)	(0.233)
	[0.000]	[0.502]	[0.000]	[0.001]	[0.000]
R^2	0.000	0.000	0.001	0.000	0.001
Obs.	9,823	9,823	9,823	9,823	9,823

Table 9. Regressions of long-short returns on the return indices of other long-short returns

This regression models the return-predictability of each characteristic, relative to its original sample period and publication date, and relative to the returns of other characteristics. The dependent variable is the monthly long-short return of a characteristic scaled, by its monthly in-sample mean. This monthly return is generated via a monthly Fama and MacBeth (1973) regression. The regression labeled *Continuous* (Panel A) uses Fama-MacBeth slopes that are generated using continuous variables. The regression labeled *Portfolios* (Panel B) uses Fama-MacBeth slopes from long-short quintile portfolios. The regression labeled *Strongest* (Panel C) uses either *Continuous* or *Portfolios* returns, depending on which method produces stronger in-sample statistical significance. *Post Publication* is equal to 1 if the month is after the official date of the journal that published the study. *Other In* is an equal weighted return of all other long-short returns for which the current month implies that the characteristic is in the original study's sample period. *Other Post* is an equal weighted return of all other long-short returns for which the current month implies that the characteristic is after the study's publication date. All indicator variables are equal to 0 if they are not equal to 1. The regressions include random effects. Standard errors are clustered on time, and are reported in parentheses. P-values are in brackets for the hypothesis that the coefficient equals 0.

Panel A: Strongest	**Coeff.**	**SE**	**P>z**
Other-In	0.634	0.03	0.000
Other-Post	0.025	0.02	0.136
*Post Pub * Other In*	-0.555	0.05	0.000
*Post Pub * Other Post*	0.399	0.06	0.000
Post Pub	-0.071	0.06	0.260
Constant	0.351	0.04	0.000
Within R^2	0.024		
Between R^2	0.126		
Overall R^2	0.024		
Obs.	30,534		

Panel B: Continuous	**Coeff.**	**SE**	**P>z**
Other-In	0.660	0.03	0.000
Other- Post	0.026	0.02	0.123
*Post Pub * Other In*	-0.596	0.05	0.000
*Post Pub * Other Post*	0.483	0.05	0.000
Post Pub	-0.093	0.06	0.121
Constant	0.317	0.04	0.000
Within R^2	0.029		
Between R^2	0.197		
Overall R^2	0.029		
Obs.	30,530		

Table 9 (Continued):

Panel C: Continuous	**Coeff.**	**SE**	**P>z**
Other-In	0.531	0.04	0.000
Other-Out	0.076	0.02	0.001
*Post Pub * Other In*	-0.424	0.06	0.000
*Post Pub * Other Post*	0.259	0.08	0.001
Post Pub	-0.112	0.09	0.202
Constant	0.453	0.06	0.000
Within R^2	0.016		
Between R^2	0.064		
Overall R^2	0.017		
Obs.	30,522		

Table 10. Autoregressions of long-short characteristic based returns.

This regression tests whether the return-predictability of the various characteristics is persistent. The dependent variable is the monthly long-short return of a characteristic scaled by its monthly in-sample mean. This monthly return is generated via a monthly Fama and MacBeth (1973) regression. We use either continuous variables or quintile portfolios to generate the characteristic's monthly return, depending on which method has stronger in-sample statistical significance. The dependent variable is regressed on lagged values, measured over the last 1, 2, and 3 months, and the average values of the lags over the last 6 and 12 months. The regressions include random effects. Standard errors are clustered on time, and are reported in parentheses. P-values are in brackets for the hypothesis that the coefficient equals 0.

	Full Sample	*In-Sample*	*Post-Publication*	*Full-Sample*	*Full-Sample*
Lag1	0.132	0.113	0.166	0.123	0.116
	(0.026)	(0.023)	(0.045)	(0.029)	(0.029)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Lag2	0.030	0.021	0.034		
	(0.024)	(0.020)	(0.042)		
	[0.210]	[0.301]	[0.420]		
Lag3	0.000	-0.014	0.033		
	(0.020)	(0.019)	(0.036)		
	[0.990]	[0.461]	[0.353]		
Sum of Lags 1-6				0.011	
				(0.010)	
				[0.271]	
Sum of Lags 1-12					0.016
					(0.006)
					[0.007]
Intercept	0.729	0.862	0.461	0.702	0.595
	(0.057)	(0.069)	(0.076)	(0.061)	(0.064)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R^2	0.019	0.014	0.033	0.019	0.023
Obs.	37,462	23,509	9,984	37,248	36,816