Open Source Cross-Sectional Asset Pricing

Andrew Y. Chen
Federal Reserve Board
andrew.y.chen@frb.gov

Tom Zimmermann
University of Cologe
tom.zimmermann@uni-koeln.de

May 2020*

Abstract

We provide data and code that successfully reproduces nearly all cross-sectional stock return predictors. Unlike most metastudies, we carefully examine the original papers to determine whether our predictability tests should produce t-stats above 1.96. For the 180 predictors that were clearly significant in the original papers, 98% of our reproductions find t-stats above 1.96. For the 30 predictors that had mixed evidence, our reproductions find t-stats of 2 on average. We include an additional 105 characteristics and 945 portfolios with alternative rebalancing frequencies to nest variables used in other metastudies. Our data covers all portfolios in Hou, Xue and Zhang (2017); 98% of the portfolios in McLean and Pontiff (2016); 90% of the characteristics from Green, Hand, and Zhang (2017); and 90% of the firm-level predictors in Harvey, Liu, and Zhu (2016) that use widely-available data.

*First posted to SSRN: May 18, 2020. We thank Ivo Welch for encouraging us to work on this project. We thank Shane Corwin, Paul Schultz, Luis Palacios, Rabih Moussawi, and Denys Glushkov for making their code publically available. The views expressed herein are those of the authors and do not necessarily reflect the position of the Board of Governors of the Federal Reserve or the Federal Reserve System.

1. Introduction

Academic finance progresses through a mixture of open collaboration and closed competition. In this paper, we attempt to push the culture toward open collaboration by providing an "open source dataset" of hundreds of predictors of the cross-section of stock returns.

Open collaboration is critical because of the rise of massive data sets and computing power. These revolutions have led to increasingly opaque analyses and a grave threat from p-hacking (a.k.a. data-snooping or data-mining). In the worst case, finance papers could become a series of unrepeatable studies, each the result of a massive and uncontrolled specification search.

Cross-sectional asset pricing is on the leading edge of these trends. The literature has reached a state where each paper studies dozens, or even hundreds of return predictors. These large datasets are often fit to non-linear models using computationally intensive algorithms (Harvey, Liu, and Zhu 2016; DeMiguel et al. 2017; Freyberger, Neuhierl, and Weber 2017; Chen and Zimmermann 2018). Working in a culture of closed competition, it is difficult (if not impossible) for new researchers to build off of these studies and further our collective understanding of risk and return.

Our open source dataset makes it easy for new researchers to contribute. Our data provides the building blocks to replicate and extend any of the aforementioned studies, and our code allows new researchers to modify and extend the building blocks themselves. Code, firm characteristics, and portfolio returns are available at https://github.com/OpenSourceAP/CrossSection. Moreover, we are committed to updating this data on an annual basis. We hope this demonstration of open collaboration will inspire others to open up their analyses.

Table 1 provides an overview of our data. Our baseline data is built from 210 firm-level return predictors that can be created from widely-available data. 180 of these are "clear predictors": the original papers clearly demonstrate that our portfolios should achieve statistical significance. Another 30 are "likely predictors": the original papers demonstrate statistical significance or strong economic significance, but our portfolios tests deviate enough from the original tests that we are unsure which side of 1.96 our t-stats will fall on. For each of predictor, we create a long-short portfolio based on the results in the original papers. This core set of predictors and portfolios forms our preferred dataset.

[Table 1: “Overview of Open Source Asset Pricing Data” around here]

We also offer an extended dataset with 315 characteristics. The 105 additional characteristics include “maybe” predictors, which were not clearly examined for predictive power, and “not predictors,” which explicitly failed to achieve significant predictability. We also include variations of characteristics formed by slightly altering the originals. We include these additional characteristics because these predictor categorizations require subjective judgments.

Unlike the baseline data, where we produce one portfolio for each characteristic, the extended dataset has four portfolios for each characteristic. The additional portfolios are formed by simply altering the rebalancing frequency. As seen in Table 1, Hou, Xue, and Zhang’s (2017) replication study features 212 portfolios of this type. If, following Hou et al., we count each of these alternatives as a distinct “anomaly,” our open source dataset offers an unprecendented library of 1,260 “anomalies.”

Our dataset is not just large—it is comprehensive. Panel B of Table 1 shows that we cover all 452 of Hou, Xue, and Zhang’s (2017) “anomalies,” 98% of the characteristics from McLean and Pontiff (2016), 92% of the characteristics from Green, Hand, and Zhang (2017), and 90% of the clear firm-level predictors that use widely-available data from Harvey, Liu, and Zhu (2016). Indeed, most of the clear, widely-available predictors that we are missing are closely related to predictors that we offer (e.g. the industry-adjusted value and momentum predictors of Asness, Porter, and Stevens 2000).

Our code reproduces predictability extremely well. We test for predictability by examining the mean returns of long-short portfolios following the original papers. 98% of the our 180 clear predictor reproductions result in t-stats above 1.96.

Much of this success is due to careful reading of the original papers. This painstaking effort was required to determine that 49 characteristics came with a lack of predictability evidence in the original papers. For example, Dimson (1979) studies a liquidity-adjusted market beta, but does not demonstrate that this adjusted beta predicts returns. As the Dimson beta was never shown to predict returns in the first place, we do not consider our statistically insignificant Dimson beta portfolio as a reproduction failure.[1]

[1] We aim for “reproductions”—that is, attempts to replicate the same result in the same sample with the same code, rather than reexaminations, to use Welch’s (2019) terminology.

This lack-of-evidence is not easy to determine, however, as Dimson (1979) does not explicitly state that it does not examine predictability. Instead, we needed to carefully comb through all of the paper's 30 pages and 10 tables to identify this lack of predictability evidence, as we did with the other 49 characteristics categorized as "maybe predictors."

Even for characteristics that were shown to be significant predictors, a careful reading is required to determine whether a particular reproduction should also produce significance. For example, Johnson and So (2012) show that option volume relative to recent averages is significant, but their test uses weekly signal updates, far more frequent than the monthly updates used in our reproduction. Moreover, the original paper's t-stat of 2.45 implies that even a tiny 10 bps change in the monthly mean return would lead to a t-stat below 1.96.[2] As a result, we categorize this and similar characteristics as "likely predictors." As one might expect, half of our likely predictors produce t-stats above 1.96. We explain in detail why each of our 30 likely predictors is categorized as such.

Despite the careful readings, we still find that four of our reproductions failed. One of these failures is the Pástor and Stambaugh (2003) liquidity beta, which is known to be fragile (Li, Robert, and Velikov 2019; Pontiff and Singla 2019). The other three failures are our reproductions of R&D productivity forecasts (Cohen, Diether, and Malloy 2013), shareholder activism (Cremers and Nair 2005), and coskewness (Harvey and Siddique 2000). Given that our study examines more than 300 characteristics, this finding should not be considered a criticism of these studies. It is quite possible that there is an error in our code. We welcome comments, and indeed contributors can directly edit the code in our repo and create a pull request at https://github.com/OpenSourceAP/CrossSection.[3]

Sections 2 and 3 describe our baseline open source dataset. There, we list the performance of each clear predictor (Table 2), each likely predictor (Table 3), and explain why each likely predictors is categorized as such (Section 3.3). Section 4 demonstrates that our baseline data displays intuitive properties. Section 5 describes the extended data. Section 6 concludes.

---

[2] To see this, note that the average standard error is 20 bps per month (McLean and Pontiff 2016; Chen and Zimmermann 2018).

[3] We use the version-control system Git to facilitate open collaboration among researchers. The system allows to track changes to our code over time, and to integrate reviewed improvements suggested by researchers into the code (pull request).

2. Baseline Characteristics Data: 210 Firm-Level Predictor Reproductions

Our predictors draw from 4 asset pricing meta-studies: McLean and Pontiff (2016); Green, Hand, and Zhang (2017); Hou, Xue, and Zhang (2017); and Harvey, Liu, and Zhu (2016). The variables studied in these papers are variously described as "characteristics," "predictors," "signals," "anomalies" or "factors." From these studies, we attempt to capture all firm-level characteristics that predict firm-level returns that can be made from widely-available data.

Our code aims for "reproductions" in Welch's (2019) terminology. That is, we attempt to replicate the same result in the same sample with the same code. This approach tests the most fundamental aspect of these predictors: whether these predictors represent repeatable events, analyzable by scientific methods (Popper 1959). Moreover, reexaminations and other reanalyses can be readily built off of our reproductions.

2.1. Definition of a Firm-Level Predictor

We define firm-level predictors using results from the original papers. These results include portfolios formed by sorting on firm characteristics, firm-level predictive regressions, and event studies, but only if these tests use only data available up to month t to predict returns in month $t+1$. We consider tests which target simple mean returns as well as characteristic- and factor-adjusted returns. In our experience, the characteristic and factor adjustments typically have little impact.

In the baseline data, we include only predictors that had evidence indicating that portfolio sorts would be significant at the 5% level. For almost all predictors, we simply require a t-stat above 1.96 in absolute value or a p-value smaller than 5%. In a couple cases, we used our judgment that the economic magnitudes are large enough to justify inclusion, despite the fact that neither a t-stat nor p-value was shown. We discuss these borderline cases in detail in Section 3.3.

This data restriction is motivated by feasibility. To make the reproduction of more than 300 characteristics achievable in a reasonable amount of economist-hours, we use the same quality check on all characteristics: we form portfolios using data available at month t and evaluate long-short returns in month $t+1$.

Thus, we exclude from the baseline data many variables that explain mean returns on size- and B/M-sorted portfolios (e.g. Lettau and Ludvigson 2001, Balvers and Huang 2007, Da 2009), as we are unsure if we should expect these variables to create statistically significant portfolio sorts. Similarly, we exclude predictors that were shown to only explain mean returns using contemporaneous data (e.g. Acharya and Pedersen 2005). We also exclude predictors that were explicitly shown to produced t-stats < 1.96 for raw return portfolio sorts (e.g. Whited and Wu 2006). Many of these additional variables are included in our extended data (Section 5).

It is impossible to exactly recreate all of these predictors. Instead, we aim to (1) capture the spirit of the original papers and (2) quantitatively match key results. In creating such a large dataset, tradeoffs must be made between comprehensiveness and faithful replication.

2.2. Standardized Constructions

We compute all characteristics at a monthly frequency. For variables that are updated at a lower frequency, the monthly value is simply the most recently observed value. This approach streamlines the code and allows for the computation of lower-frequency versions based on the monthly data. However, our data constructions may deviate from the original papers, which sometimes compute only annual statistical tests.

We assume the standard six-month lag for annual accounting data availability and a one-quarter lag for quarterly accounting data availability. For IBES, we assume earnings estimates are available by the statistical period end date. Other data is assumed to be available following the original papers.

Many characteristics were only shown to be predictive in particular subsets of the data. We try to put off subsetting until the portfolio generation step. Thus, the characteristics code and data omits price and exchange filters, which are instead imposed in portfolio generation (Section 3).

Other filters, however, are quite diverse and difficult to implement at the portfolio stage. Several papers exclude stocks based on SIC codes or missing accounting data. Still others find predictability only in value stocks (Piotroski 2000) or find that predictability only exists in subsets of the predictor itself (Dichev 1998). To accommodate these filters in a manageable fashion, we set to missing firm-months that don't satisfy these filters in the characteristics code.

2.3. Distinct Predictors

We make no attempt to eliminate predictors due to subjective similarities. Thus, we include several profitability-related predictors including those from Fama and French (2006); Balakrishnan, Bartov, and Faurel (2010); and Novy-Marx (2013). Being liberal about distinct predictors is necessary as there is, as of yet, no established methodology for determining distinct predictors. By including all predictors, we allow future users of our code and data to make their own determination on which version of profitability is the "right" one.

Despite this potential redundancy, a simple analysis suggests that this dataset is very high-dimensional. Figure 1 shows that the distribution of pairwise rank correlations between predictors is centered around zero (Panel (a)), and indeed 90% of correlations are less than 0.25 in absolute value. This near-zero correlation exists despite the fact that we sign all predictors so that a larger value implies a higher expected return.

[Figure 1 "Pairwise Rank Correlations Between Firm-Level Predictors"about here.]

Panels (b)-(d) provides some more detail, showing that the 200+ predictors are largely distinct from the prominent predictors like B/M, momentum (12-month), and gross profitability. These results are consistent with Green, Hand, and Zhang (2013) who also find correlations close to zero among their set of 39 readily programmed predictors.

3. Baseline Portfolios: 210 Long-Short Strategies

For many applications, researchers do not need to dig into the firm level data as they only require portfolio returns. Thus, we offer a set of 210 long-short portfolio returns, one from each baseline characteristic. Portfolios are implemented following results in the original papers. This approach is also used in McLean and Pontiff (2016) and aims for simplicity and flexibility. It also allows for a simple check of the quality of our characteristic reproductions.

Thus, most portfolios use equal-weighted quintiles, as nearly all papers use either equal-weighted portfolio sorts or equal-weighted regressions. We use value-weighting or other quantiles for the handful of papers that emphasize these constructions. We also rebalance following the original papers or when the predictor updates (annually for annual Compustat variables). Price and exchange filters are also chosen following the original papers.

Figure 2 shows that our data produces many distinct portfolios. The figure plots the distribution of pairwise correlations between portfolio returns. Portfolios are signed to all have positive mean returns, and yet the mean, median, and modal correlation is very close to zero. Indeed the vast majority of correlations lie between -0.5 and +0.5.

[Figure 2 "pairwise correlations between predictor portfolio returns" about here.]

3.1. Categorization of Clear vs Likely Predictors

We categorize predictors as "clear" or "likely" based on the results in the original papers. Clear predictors are those that we expect to achieve statistical significance in our portfolio sorts. Likely predictors were shown to be strong in the original papers, but we are unsure about what to expect in our tests.

This categorization is necessary because our portfolio sort t-stats are a blend of "reproductions" and "reexaminations" (Welch 2019). Though we try to be faithful as possible to the original papers (reproductions), for practical reasons both our characteristic constructions and portfolio tests deviate from the originals (reexaminations). As a result, in some cases we are unsure if we should expect to achieve significance, despite the significant results in the original papers.

For 86% of the 210 firm-level predictors, we find that we should expect significance in our tests based on the results in the original papers. These clear predictors are typically easy to classify, as many papers show portfolio sorts, produce very large t-stats, and use simple combinations of accounting and market price data that we can reproduce closely. In other cases, the original papers only show regressions or event studies, but the significance is so strong that we expect to obtain significance in our portfolios.

For the remaining 30 predictors, the typical case is that the original test showed only marginal significance in a test that is different than ours. In other cases, the original paper found strong predictability, but only in tests that are markedly different than ours.

We discuss our predictor categorizations in excruciating detail in Section 3.3, where we discuss the performance of individual likely predictors.

3.2. Performance of Clear Predictor Reproductions

To measure the quality of our clear predictors, we examine the t-stat for the mean return on our long-short portfolios using the original papers' sample periods. Following our definition of a clear predictor (Sections 2.1 and 3.1), we judge t-stats > 1.96 as reproduction successes.

Our reproductions are extremely successful. 98% of clear predictor reproductions have t-stats above 1.96. Figure 3 breaks down the success rate by broad data categories. 130 of the clear predictors focus on accounting or stock price data, and our success rates are effectively 100% for these major data categories. The remaining 50 use a wide variety of data sources: 13F institutional holdings, analyst forecasts, events, trading data, and as well as "other" data. Across these other categories, our reproduction success rates remain close to 100%.

[Figure 3 "Reproduction Success Rates for Clear Predictors" about here.]

Indeed, only one data category leads to reproduction rates below 98%. This category is 13F data, with its reproduction rate of 86%. This 14% failure rate, however, represents only a single predictor, as only 7 of our clear predictors use 13F data.

For a closer look, Table 2 lists the performance of all 180 of our clear predictors. Predictors are sorted by author names, so the reader can easily browse our dataset and check the performance of particular reproductions.

[Table 2 "Predictor-Level Performance" about here.]

In what follows, we discuss failed reproductions, selected marginal successes, and selected predictors with extremely large t-stats.

3.2.1. Failed Reproductions

Only four of our 180 clear predictors produce t-stats < 1.96. One is the Pástor and Stambaugh (2003) beta, which is known to be fragile (Li, Robert, and Velikov 2019; Pontiff and Singla 2019). The other three are R&D productivity forecasts (Cohen, Diether, and Malloy 2013), governance-among-blockheld (Cremers and Nair 2005) and coskewness (Harvey and Siddique 2000). These reproductions were complicated by a variety of issues that may have led to our low t-stats.

Our reproduction of Cohen, Diether, and Malloy's (2013) R&D productivity forecast portfolios achieves a low t-stat of 0.73. These forecasts are created by rolling estimation of a sales forecasting model that involves several lags of R&D. As R&D is prone to missing and zero values, it is quite possible that we failed to follow the exact same procedures as the original authors.

Our reproduction of Cremers and Nair's (2005) governance-among-blockheld portfolio generates a t-stat of 1.75. This portfolio combines data from two specialized sources: 13F institutional holdings and the Gompers, Ishii, and Metrick (2003) governance index. Neither of us is an expert in either topic. Indeed, we had some difficulty constructing our own institutional ownership variables from the raw s34 types 1 and 3 datasets, and ended up using code provided by WRDS.[4]

Our reproduction of Harvey and Siddique's (2000) coskewness portfolios generates a t-stat of 0.32. Coskewness is constructed from rolling computation of the sample coskewness between the stock's return and the market's using monthly CRSP data. This clear predictor is unique in that the original paper does not present tables or figures on this portfolio. Instead, the authors simply state in the text that they "reject the hypothesis that the mean spread is zero at the 5 percent level." The combination of a rolling estimation and absence of tables on the portfolios could easily lead to our code deviating significantly from that of the original authors'. Indeed, categorizing coskewness as a clear predictor is a judgment call, and it could easily be considered a likely predictor.

As emphasized in the introduction, our reproductions failures should not be taken as a criticism of these papers. In reproducing hundreds of characteristics, it is quite possible that there is an error in our code. We would be grateful to readers who notify us of errors and will update the open source data accordingly.

[4] We are grateful to Luis Palacios, Rabih Moussawi, and Denys Glushkov for making their code available.

3.2.2. Marginally Significant Reproductions

27 of our reproductions produce t-stats between 1.96 and 2.50. Several of these marginally significant t-stats come from our implementations of event studies. These predictors include the dividend omission predictor from Michaely, Thaler, and Womack (1995); the recent IPO indicator from Ritter (1991); and the credit rating downgrade indicator from Dichev and Piotroski (2001). This borderline performance is intuitive given that these papers were event studies, while our statistical evaluation uses portfolio sorts. Indeed, some readers may wish to exclude event studies from our list of clear predictors.

Other marginal reproductions come from tests that deviate significantly from the original papers. Our reproduction of the industry-linkage portfolios of Menzly and Ozbas (2010) produce t-stats of 2.04 and 2.07, but our portfolios are produced by sorting individual stocks, while Menzly and Ozbas sort industry portfolios. Indeed, in unreported results, we find that sorting industry portfolios leads to much larger t-stats. Our option volume to stock volume portfolios (Johnson and So 2012) lead to a t-stat of 2.07, but this result comes from monthly signal updates, while the original paper used weekly signal updates. Both of these deviations come from the standardization of portfolio implementations that is required for replicating hundreds of portfolios.

Finally, some marginal t-stats come from marginal results in the original papers. Hou and Robinson (2006) find that the raw return of their baseline industry concentration measure produces a t-stat of 2.14, not far from our t-stat of 2.33. Similarly, Palazzo (2012) finds that cash to assets produces a raw return t-stat 2.14, very close to our t-stat of 1.99. Indeed, while we judged these predictors to be "clear" as they provided raw return portfolio sort t-stats, some readers may consider these predictors to be "likely," as our code most certainly deviates from the original papers.[5]

3.2.3. Extremely Significant Reproductions

Moving to the other side of the significance spectrum, 46 clear predictor reproductions found enormous t-stats in excess of 5.0. To understand the magnitude of this t-stat, the corresponding p-value is 0.0000003. It is absurdly unlikely that these predictors are drawn from the null of no predictability.

[5] Both of these papers provide substantial evidence for the relevance of their predictors beyond the raw return t-stats, and indeed their other results tend to be much stronger.

Almost all of these outstanding predictors focus on accounting data, analyst forecasts, or stock prices. Stated differently, almost none of them come from the more exotic data categories. This extreme performance may reflect the higher quality of the more common data sources, or possibly lower performance standards required to publish novel data.

These outstanding performers are quite diverse. They include consecutive earnings increases (Loh and Warachka 2012); net external financing (Bradshaw, Richardson, and Sloan 2006), change in recommendation (Jegadeesh et al. 2004), return skewness (Bali, Engle, and Murray 2016), return seasonality (Heston and Sadka 2008), conglomerate return (Cohen and Lou 2012), dividend indicator (Hartzmark and Solomon 2013), employment growth (Belo, Lin, and Bazdresch 2014), asset growth (Cooper, Gulen, and Schill 2008), the Kaplan-Zingales index (Lamont, Polk, and Saá-Requejo 2001), abnormal accruals (Xie 2001), change in inventory (Thomas and Zhang 2002), and enterprise multiple (Loughran and Wellman 2011). These predictors lack any obvious economic connection, consistent with the near zero median correlation in Figure 2.

3.3. Performance of Individual Likely Predictors

Evaluating the quality of our likely predictors is more difficult. Likely predictors, by definition, are predictors where we are unsure of which side of 1.96 our portfolio t-stats will end up on, based on the results in the original papers. Thus, Table 3 simply lists all 30 likely predictors along with their mean in-sample returns and t-stats. The table shows performance in-line with what one might expect given the definition of a likely predictor: the median and mean of these t-stats are both close to 2.

[Table 3 “Performance of Individual Likely Predictors” about here.]

The table also helps illustrate why we judge predictors to be “likely.” At the top of the table we have sales growth over inventory growth, from Abarbanell and Bushee (1998). Abarbanell and Bushee did not show portfolio sorts, and showed forecasting regressions with several independent variables. They found that this characteristic’s coefficient had a t-stat of 2.06. Regressions tend to generate larger t-stats than our portfolio sorts (be more powerful), but including multiple independent variables tends to reduce t-stats. Thus, we are unsure if our portfolio sorts should produce a t-stat larger or smaller than 1.96. As seen in Table 3, our reproduction turns out to have a t-stat of 2.83, but this significance could not be known from Abarbanell and Bushee's (1998) regression results.

Many of the likely predictors are similar to sales growth over inventory growth: the original papers showed marginal significance using statistical tests that are different than ours. Some were tested with regressions: sales growth over overhead growth (Abarbanell and Bushee 1998), sales to price (Barbee, Mukherji, and Raines 1996, beta (Fama and MacBeth 1973), change in asset turnover (Soliman 2008), secured debt (Valta 2016), deferred revenue (Prakash and Sinha 2013). Others were tested in event studies: spinoffs (Cusatis, Miles, and Woolridge 1993), dividend initiation (Michaely, Thaler, and Womack 1995). Still others used long-short portfolio returns with factor or characteristic adjustments: brand investment (Belo, Lin, and Vitorino 2011), industry concentration based on assets or equity (Hou and Robinson 2006). For all of these predictors, the original papers found t-stats between 1.96 and 2.60. As seen in Table 3, our reproductions largely lead to t-stats in this range. Only the secured debt predictors and cash-based operating profitability perform notably worse (t-stats of 0.99-1.73), and only change in asset turnover performs better (t-stat of 4.2).

For other likely predictors, the original papers presented strong predictability, but our reproductions deviate moderately from the original tests. Our earnings forecast to price variable uses all stocks, while the predictability found in Elgers, Lo, and Pfeiffer (2001) came largely from low analyst coverage stocks. Similarly moderate deviations from the original tests are found in our reproductions of pension funding status (Franzoni and Marin 2006), operating profitability (Fama and French 2006), and growth in net long term operating assets (Fairfield, Whisenant, and Yohn 2003). As it turns out, all of these predictors produced t-stats above 1.96 in our reproductions.

In other cases, our reproductions deviate markedly. Amihud and Mendelson (1986) used quoted bid-ask spreads from Fitch's Stock Quotations on the NYSE to predict returns, but our construction uses Corwin and Schultz's (2012) effective bid-ask spread proxy based on daily CRSP prices (following McLean and Pontiff 2016). Similarly, Frazzini and Pedersen's (2014) betting-against-beta factor used a non-standard construction where each stock is weighted depending on its beta ranking, far from our simple equal-weighted construction. Our adjusted price delay predictor (Hou and Moskowitz 2005) comes from rolling regressions of daily individual stock returns, while the original paper conducted a 2-stage procedure that first estimates a noisy measure of price delay on individual stocks and then reduces the noise by running a second set of regressions on portfolios formed from the first stage. Other predictors in this group include analyst value and analyst optimism (Frankel and Lee 1998), sin stocks based on selection criteria (Hong and Kacperczyk 2009), option volume relative to recent averages (Johnson and So 2012), and volatility smirk (Xing, Zhang, and Zhao 2010) . As expected from reproductions that deviate from the originals, our portfolios generate a wide range of t-stats, ranging from 0.27 (for the price delay adjusted for standard errors) to 5.31 (for analyst optimism).

For further details on our predictor categorization judgments, see SignalDocumentation.xlsx at https://github.com/OpenSourceAP/CrossSection. We welcome pull requests from readers who are interested in helping us more closely match the original papers.

4. Kicking the Tires: Portfolio Performance by Rebalancing Frequency and Liquidity Screen

This section shows that portfolio performance declines if we either (1) increase the rebalancing frequency or (2) impose liquidity screens. These intuitive results further demonstrate the quality of our dataset. They also demonstrate the flexibility of our open source data.

4.1. Performance by Rebalancing Frequency

Our code allows for a flexible choice of the rebalancing frequency, irrespective of the frequency of the raw characteristic. More precisely, the code allows the user to choose how often the portfolio should take on new characteristics values in deciding which stocks to have in the long and short legs. Given the characteristics in use, stock weights are rebalanced every month to ensure equal- or value-weighting, following the predictability literature.[6] This flexibility may be important, for example, when accounting for trading costs.

Figure 4 shows that our code produces intuitive results when we alter the rebalancing frequency. This figure plots the distribution of mean returns across

[6]Most papers do provide precise explanations of rebalancing, but in our experience this procedure is required for replicating papers. For an explicit example, see https://wrds-www.wharton.upenn.edu/pages/support/applications/risk-factors-and-industry-benchmarks/fama-french-factors/.

predictors for 1-, 3-, 6-, and 12-month rebalancing. For clarity, we use only predictors that use 1-month rebalancing in the baseline.

[Figure 4 "Performance by Rebalancing Frequency" about here.]

The figure shows that performance declines monotonically as the rebalancing period increases from 1- to 12-months. The median predictor's mean return declines from about 65 to 40 bps per month. The decline for 75th percentile predictor is even sharper, falling from about 105 to 65 bps.

These results are intuitive: less frequent rebalancing implies less exposure to the predictive signal. Alternatively, strategies that earn higher gross returns entail higher trading costs.

4.2. Performance by Liquidity Screen

Our code also flexibly implements liquidity screens. Many of the original papers apply liquidity screens, and we try to follow these screens in our baseline reproductions. however, these screens are imposed at the portfolio generation step rather than at the characteristic step.

Figure 5 shows that our code produces intuitive liquidity effects. The figure shows the distribution of mean returns after imposing various liquidity screens. The entire distribution of shifts toward zero on the imposition of any screen.

[Figure 5 "Performance by Liquidity Screen" about here.]

The price screen (limiting to stocks with share price > $5) appears to be the softest, as it produces the smallest decline in mean returns. The median predictor's mean return declines from about 60 bps per month before the screen to about 45 bps per month after the screen. The NYSE screen (limiting positions to NYSE stocks) results in a sharper decline and a median mean return of 35 bps per month. The market equity screen (limiting to stocks with market equity > the 20th percentile among NYSE stocks) leads to similar results as the NYSE screen.

5. Extended Dataset

Our baseline data contains only the 210 original predictors that were shown to predict firm-level returns in the original papers. Some researchers may be interested in a broader set of characteristics and portfolios, however. For these researchers, we offer an extended dataset of 315 characteristics and 1260 portfolios.

Characteristics in the extended dataset fall into two broad categories: original characteristics and variants. Original characteristics can be traced to results in the original papers. Section 5.1 discusses these original characteristics as well as the performance of portfolios built off of these characteristics. There we explain how assign characteristics to the "maybe" and "not" predictor categories.

Variants are formed by arbitrary modifications of the original characteristics. Similarly, the extended data also contains many portfolios that only change the rebalancing frequency of other portfolios. These variables are included solely to nest the Hou, Xue, and Zhang (2017) replication study. We briefly discuss these variations in Section 5.2.

5.1. Additional Original Characteristics

Our extended dataset contains 59 original characteristics (not variants) beyond those in the baseline data. We categorize these characteristics into two groups: maybe predictors and not predictors. For the 49 maybe predictors, the original papers did not present clear predictability evidence. The 10 not predictors were explicitly shown to fail to produce statistically significant predictability in the original papers.

Figure 6 shows that, unlike the clear and likely predictors, predictors in the maybe and not categories generally failed to achieve statistical significance in our reproductions. Maybe predictors generally performed better than not predictors, and only one of our not predictors "failed" in the sense that we produced a t-stat *higher* than 1.96.

[Figure 6 "Predictive Significance in the Extended Dataset" about here.]

For a closer look, Table 4 lists every additional characteristic along with their predictor categories and in-sample performance. The table is grouped by predictor category to aid in understanding our categorizations. Within predictor category, we sort characteristics based on the authors of the original papers for ease of reference.

[Table 4 "Additional Characteristics" about here.]

In the remainder of this subsection, we discuss selected characteristics within each predictor category.

5.1.1. Selected Maybe Predictors

Most of the maybe predictors simply did not come with predictability evidence in the original papers. Several of these "no evidence" predictors study the implied cost of capital. For example, accrual quality, earnings conservatism, and earnings value relevance all come from Francis et al. (2004), which studies characteristics that are related to an implied cost of capital estimate based on Value Line's price targets. This paper does not, however, examine return prediction. Similarly, two of the maybe predictors come from Ortiz-Molina and Phillips (2014), which examines the relationship between an implied cost-of-capital estimate and these measures of asset liquidity.[7] It's worth noting that the implied cost of capital literature often emphasizes that the costs of capital may differ from the mean of realized returns (Gebhardt, Lee, and Swaminathan 2001).

Other "no evidence" predictors include the number of analysts from Elgers, Lo, and Pfeiffer (2001) and the brand capital measure from Belo, Lin, and Bazdresch (2014). These papers did examine predictability—indeed our reproductions of their predictive characteristics succeed—but they did not examine predictability for the aforementioned characteristics. The number of analysts was used as a control variable in Elgers et al, and brand capital was simply a component of the model in Belo et al. Similarly, the Dimson's (1979) beta was simply not examined for predictability.

Table 4 shows that the aforementioned "no evidence" predictors generally produce insignificant t-stats in our reproductions. The exception is asset liquidity scaled by market value of assets which, like other scaled-price measures, produces highly significant predictability.

Other maybe predictors came with predictability-related information in the original papers, but we judged this evidence as too weak to be considered relevant for judging statistically significant predictability in our portfolio sorts. Several of these weak evidence predictors come from Acharya and Pedersen's (2005) study of liquidity betas. Acharya and Pedersen estimated market prices of risk for these betas in a GMM framework, which would imply predictability if the pa-

[7] Ortiz-Molina and Phillips do demonstrate predictability for measures of potential mergers using data on rival firms, but we did not attempt to reproduce these predictors.

rameters are very stable. But since betas tend to be unstable (Ang, Chen, and Xing 2006, for example) we judge this GMM result as close to no evidence regarding the results of portfolio sorts. As it turns out, Table 4 shows that none of our reproductions of these liquidity betas generate t-stats above 1.96.

Other predictors that came with very weak predictability-related evidence come from three papers that fit large models to firm-level data: Ou and Penman (1989); Haugen and Baker (1996); and Holthausen and Larcker (1992). These papers find that their fitted models predict returns, but we do not reproduce their models. Instead, we simply examine the inputs in these models, following Green, Hand, and Zhang (2017). We judge the fitted model results as too distant to our individual characteristics to be considered evidence of predictability, but we acknowledge that this is a judgment call. Indeed, Table 4 shows that the performance of the individual characteristics is respectable, with t-stats averaging around 1.5.

Last, we include as maybe predictors characteristics that were studied in multivariate regressions, but were shown to be statistically insignificant. This judgment of "no evidence" is found by process of elimination. We certainly cannot consider these results as indicating that our single characteristic portfolio sorts should also generate insignificance, as the regression results can depend strongly on the controls. At the same time, we cannot consider insignificant regression coefficients as indicating significant portfolio sorts. These predictors include several from Soliman (2008) (change in noncurrent operating assets, change in noncurrent operating liabilities, profit margin, and return on net operating assets), as well as a few from Abarbanell and Bushee (1998) (labor force efficiency, effective tax rate, sales growth to receivables growth, and change in growth margin vs sales). As seen in Table 4, most of our portfolio implementations of these characteristics lead to insignificant t-stats.

5.1.2. Selected Not Predictors

The extended data includes only a handful of not predictors. These not-predictors come from an assortment of papers and are difficult to classify. The Whited and Wu (2006) financial constraints index was shown to be almost-but-not-quite significant in the original paper. The realized downside beta of Ang, Chen, and Xing (2006), was shown explicitly to fail to predict returns, consistent with their finding that past downside beta does not predict future downside beta. The R&D to sales ratio from Chan, Lakonishok, and Sougiannis (2001), was shown to fail to predict returns, but the same paper showed that both advertising to market cap and R&D to market cap predict returns.

As seen in Table 4, we successfully reproduce all but one of these not predictors. As in the original papers, all but one of our reproductions produces a t-stat below 1.96. The only exception is the returns not in the same month over the past 11 to 15 years characteristic (Heston and Sadka 2008), which in our reproduction finds a t-stat of 2.11. The original paper found a t-stat of 1.77, which is not far from our result. Indeed, one may want to generalize our notion of a "likely firm-level predictor" to include predictors that originally had t-stats just below 1.96.

It's important to mention that the lack of significance of these predictors should not be considered a criticism of the original papers. The 5% significance cutoff is arbitrary, and some of these predictors fall just below the cutoff. Moreover, the economics in these papers never boils down to just the statistical significance of predictor-portfolios. For example, Campbell, Hilscher, and Szilagyi (2008) show that their benchmark distress risk portfolios generate raw t-stats of 1.41, but the mean return has the opposite sign of that implied by theory. Moreover, we find that a monthly version of their distress risk portfolios generates a highly significant t-stat of 3.42.

5.2. Characteristic Variants and Portfolios with Alternative Rebalancing Frequencies

To nest Hou, Xue, and Zhang (2017), we provide many more variables that are of little interest for future researchers. These additional variables include 43 characteristics that are formed by modifying characteristics in the original papers. Most of these modifications use quarterly versions of annual accounting variables. A few involve arbitrary lags of the denominator or using alternative factor model adjustments when generating return residuals (as in idiosyncratic volatility). We also include 4 portfolios for each characteristic. These portfolios are formed by rebalancing at the 1-, 3-, 6-, and 12-month horizons.

6. Conclusion

We push the asset pricing literature toward open collaboration by providing an open source dataset that successfully reproduces nearly all cross-sectional stock return predictors. We hope that many future studies take advantage of our efforts and build off our dataset.

More importantly, we hope that future researchers take our approach as an example, and open their analyses to the world. We believe such a shift in culture is important for ensuring that academic research continues to contribute to our collective understanding of finance.

References

Abarbanell, Jeffery S and Brian J Bushee. "Abnormal returns to a fundamental analysis strategy". *Accounting Review* (1998), pp. 19–45.

Acharya, Viral V and Lasse Heje Pedersen. "Asset pricing with liquidity risk". *Journal of financial Economics* 77.2 (2005), pp. 375–410.

Amihud, Yakov and Haim Mendelson. "Asset pricing and the bid-ask spread". *Journal of financial Economics* 17.2 (1986), pp. 223–249.

Ang, Andrew, Joseph Chen, and Yuhang Xing. "Downside risk". *The review of financial studies* 19.4 (2006), pp. 1191–1239.

Asness, Clifford S, R Burt Porter, and Ross L Stevens. "Predicting stock returns using industry-relative firm characteristics". *Available at SSRN 213872* (2000).

Balakrishnan, Karthik, Eli Bartov, and Lucile Faurel. "Post loss/profit announcement drift". *Journal of Accounting and Economics* 50.1 (2010), pp. 20–41.

Bali, Turan G, Robert F Engle, and Scott Murray. *Empirical asset pricing: The cross section of stock returns.* John Wiley & Sons, 2016.

Balvers, Ronald J and Dayong Huang. "Productivity-based asset pricing: Theory and evidence". *Journal of Financial Economics* 86.2 (2007), pp. 405–445.

Barbee, William C, Sandip Mukherji, and Gary A Raines. "Do sales–price and debt–equity explain stock returns better than book–market and firm size?" *Financial Analysts Journal* 52.2 (1996), pp. 56–60.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch. "Labor hiring, investment, and stock return predictability in the cross section". *Journal of Political Economy* 122.1 (2014), pp. 129–177.

Belo, Frederico, Xiaoji Lin, and Maria Ana Vitorino. "Advertising, Brand Capital, and Asset Returns" (2011).

Bradshaw, Mark T, Scott A Richardson, and Richard G Sloan. "The relation between corporate financing activities, analysts' forecasts and stock returns". *Journal of Accounting and Economics* 42.1-2 (2006), pp. 53–85.

Campbell, John Y, Jens Hilscher, and Jan Szilagyi. "In search of distress risk". *The Journal of Finance* 63.6 (2008), pp. 2899–2939.

Chan, Louis KC, Josef Lakonishok, and Theodore Sougiannis. "The stock market valuation of research and development expenditures". *The Journal of Finance* 56.6 (2001), pp. 2431–2456.

Chen, Andrew Y and Tom Zimmermann. “Publication Bias and the Cross-Section of Stock Returns”. *Available at SSRN: https://ssrn.com/abstract=2802357* (2018).

Cohen, Lauren, Karl Diether, and Christopher Malloy. “Misvaluing innovation”. *The Review of Financial Studies* 26.3 (2013), pp. 635–666.

Cohen, Lauren and Dong Lou. “Complicated firms”. *Journal of financial economics* 104.2 (2012), pp. 383–400.

Cooper, Michael J, Huseyin Gulen, and Michael J Schill. “Asset growth and the cross-section of stock returns”. *the Journal of Finance* 63.4 (2008), pp. 1609–1651.

Corwin, Shane A and Paul Schultz. “A simple way to estimate bid-ask spreads from daily high and low prices”. *The Journal of Finance* 67.2 (2012), pp. 719–760.

Cremers, KJ Martijn and Vinay B Nair. “Governance mechanisms and equity prices”. *the Journal of Finance* 60.6 (2005), pp. 2859–2894.

Cusatis, Patrick J, James A Miles, and J Randall Woolridge. “Restructuring through spinoffs: The stock market evidence”. *Journal of Financial Economics* 33.3 (1993), pp. 293–311.

Da, Zhi. “Cash flow, consumption risk, and the cross-section of stock returns”. *The Journal of Finance* 64.2 (2009), pp. 923–956.

DeMiguel, Victor, Alberto Martin-Utrera, Francisco J Nogales, and Raman Uppal. “A portfolio perspective on the multitude of firm characteristics” (2017).

Dichev, Ilia D. “Is the risk of bankruptcy a systematic risk?” *the Journal of Finance* 53.3 (1998), pp. 1131–1147.

Dichev, Ilia D and Joseph D Piotroski. “The long-run stock returns following bond ratings changes”. *The Journal of Finance* 56.1 (2001), pp. 173–203.

Dimson, Elroy. “Risk measurement when shares are subject to infrequent trading”. *Journal of Financial Economics* 7.2 (1979), pp. 197–226.

Elgers, Pieter T, May H Lo, and Ray J Pfeiffer. “Delayed security price adjustments to financial analysts’ forecasts of annual earnings”. *The Accounting Review* 76.4 (2001), pp. 613–632.

Fairfield, Patricia M, J Scott Whisenant, and Teri Lombardi Yohn. “Accrued earnings and growth: Implications for future profitability and market mispricing”. *The accounting review* 78.1 (2003), pp. 353–371.

Fama, Eugene F and Kenneth R French. “Profitability, investment and average returns”. *Journal of financial economics* 82.3 (2006), pp. 491–518.

Fama, Eugene F and James D MacBeth. "Risk, return, and equilibrium: Empirical tests". *The Journal of Political Economy* (1973), pp. 607–636.

Francis, Jennifer, Ryan LaFond, Per M Olsson, and Katherine Schipper. "Costs of equity and earnings attributes". *The accounting review* 79.4 (2004), pp. 967–1010.

Frankel, Richard and Charles MC Lee. "Accounting valuation, market expectation, and cross-sectional stock returns". *Journal of Accounting and economics* 25.3 (1998), pp. 283–319.

Franzoni, Francesco and Jose M Marin. "Pension plan funding and stock market efficiency". *the Journal of Finance* 61.2 (2006), pp. 921–956.

Frazzini, Andrea and Lasse Heje Pedersen. "Betting against beta". *Journal of Financial Economics* 111.1 (2014), pp. 1–25.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber. *Dissecting characteristics nonparametrically*. Tech. rep. National Bureau of Economic Research, 2017.

Gebhardt, William R, Charles MC Lee, and Bhaskaran Swaminathan. "Toward an implied cost of capital". *Journal of accounting research* 39.1 (2001), pp. 135–176.

Gompers, Paul, Joy Ishii, and Andrew Metrick. "Corporate governance and equity prices". *The quarterly journal of economics* 118.1 (2003), pp. 107–156.

Green, Jeremiah, John RM Hand, and X Frank Zhang. "The characteristics that provide independent information about average us monthly stock returns". *The Review of Financial Studies* (2017), hhx019.

— "The supraview of return predictive signals". *Review of Accounting Studies* 18.3 (2013), pp. 692–730.

Hartzmark, Samuel M and David H Solomon. "The dividend month premium". *Journal of Financial Economics* 109.3 (2013), pp. 640–660.

Harvey, Campbell R, Yan Liu, and Heqing Zhu. "... and the cross-section of expected returns". *The Review of Financial Studies* 29.1 (2016), pp. 5–68.

Harvey, Campbell R and Akhtar Siddique. "Conditional skewness in asset pricing tests". *The Journal of Finance* 55.3 (2000), pp. 1263–1295.

Haugen, Robert A and Nardin L Baker. "Commonality in the determinants of expected stock returns". *Journal of Financial Economics* 41.3 (1996), pp. 401–439.

Heston, Steven L and Ronnie Sadka. "Seasonality in the cross-section of stock returns". *Journal of Financial Economics* 87.2 (2008), pp. 418–445.

Holthausen, Robert W and David F Larcker. "The prediction of stock returns using financial statement information". *Journal of accounting and economics* 15.2-3 (1992), pp. 373–411.

Hong, Harrison and Marcin Kacperczyk. "The price of sin: The effects of social norms on markets". *Journal of Financial Economics* 93.1 (2009), pp. 15–36.

Hou, Kewei and Tobias J Moskowitz. "Market frictions, price delay, and the cross-section of expected returns". *The Review of Financial Studies* 18.3 (2005), pp. 981–1020.

Hou, Kewei and David T Robinson. "Industry concentration and average stock returns". *The Journal of Finance* 61.4 (2006), pp. 1927–1956.

Hou, Kewei, Chen Xue, and Lu Zhang. *Replicating Anomalies.* Tech. rep. National Bureau of Economic Research, 2017.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D Krische, and Charles MC Lee. "Analyzing the analysts: When do recommendations add value?" *The journal of finance* 59.3 (2004), pp. 1083–1124.

Johnson, Travis L and Eric C So. "The option to stock volume ratio and future returns". *Journal of Financial Economics* 106.2 (2012), pp. 262–286.

Lamont, Owen, Christopher Polk, and Jesús Saá-Requejo. "Financial constraints and stock returns". *The review of financial studies* 14.2 (2001), pp. 529–554.

Lettau, M. and S. Ludvigson. "Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying". *Journal of Political Economy* (2001), pp. 1238–1287.

Li, Hongtao, Novy-Marx Robert, and Mihail Velikov. "Liquidity Risk and Asset Pricing". *Critical Finance Review* 8.1-2 (2019), pp. 223–255.

Loh, Roger K and Mitch Warachka. "Streaks in earnings surprises and the cross-section of stock returns". *Management Science* 58.7 (2012), pp. 1305–1321.

Loughran, Tim and Jay W Wellman. "New evidence on the relation between the enterprise multiple and average stock returns". *Journal of Financial and Quantitative Analysis* 46.6 (2011), pp. 1629–1650.

McLean, R David and Jeffrey Pontiff. "Does academic research destroy stock return predictability?" *The Journal of Finance* 71.1 (2016), pp. 5–32.

Menzly, Lior and Oguzhan Ozbas. "Market segmentation and cross-predictability of returns". *The Journal of Finance* 65.4 (2010), pp. 1555–1580.

Michaely, Roni, Richard H Thaler, and Kent L Womack. "Price reactions to dividend initiations and omissions: Overreaction or drift?" *the Journal of Finance* 50.2 (1995), pp. 573–608.

Novy-Marx, Robert. “The other side of value: The gross profitability premium”. *Journal of Financial Economics* 108.1 (2013), pp. 1–28.

Ortiz-Molina, Hernán and Gordon M Phillips. “Real asset illiquidity and the cost of capital”. *Journal of Financial and Quantitative Analysis* 49.1 (2014), pp. 1–32.

Ou, Jane A and Stephen H Penman. “Financial statement analysis and the prediction of stock returns”. *Journal of accounting and economics* 11.4 (1989), pp. 295–329.

Palazzo, Berardino. “Cash holdings, risk, and expected returns”. *Journal of Financial Economics* 104.1 (2012), pp. 162–185.

Pástor, L’uboš and Robert F Stambaugh. “Liquidity risk and expected stock returns”. *Journal of Political economy* 111.3 (2003), pp. 642–685.

Piotroski, Joseph D. “Value investing: The use of historical financial statement information to separate winners from losers”. *Journal of Accounting Research* (2000), pp. 1–41.

Pontiff, Jeffrey and Rohit Singla. “Liquidity Risk?” *Critical Finance Review* 8.1-2 (2019), pp. 257–276.

Popper, Karl R. “The logic of scientific discovery (1st English ed.)” *London: Hutchinson & Co. First published (1935), Logik der Forschung, Vienna: Verlag von Julius Springer* (1959).

Prakash, Rachna and Nishi Sinha. “Deferred revenues and the matching of revenues and expenses”. *Contemporary Accounting Research* 30.2 (2013), pp. 517–548.

Ritter, Jay R. “The long-run performance of initial public offerings”. *The journal of finance* 46.1 (1991), pp. 3–27.

Soliman, Mark T. “The use of DuPont analysis by market participants”. *The Accounting Review* 83.3 (2008), pp. 823–853.

Thomas, Jacob K and Huai Zhang. “Inventory changes and future returns”. *Review of Accounting Studies* 7.2-3 (2002), pp. 163–187.

Valta, Philip. “Strategic default, debt structure, and stock returns”. *Journal of Financial and Quantitative Analysis* 51.1 (2016), pp. 197–229.

Welch, Ivo. “Reproducing, Extending, Updating, Replicating, Reexamining, and Reconciling”. *Critical Finance Review* 8.1-2 (2019), pp. 301–304. ISSN: 2164-5744. URL: `http://dx.doi.org/10.1561/104.00000082`.

Whited, Toni M and Guojun Wu. “Financial constraints risk”. *The Review of Financial Studies* 19.2 (2006), pp. 531–559.

Xie, Hong. “The mispricing of abnormal accruals”. *The accounting review* 76.3 (2001), pp. 357–373.

Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao. “What does the individual option volatility smirk tell us about future equity returns?” *Journal of Financial and Quantitative Analysis* 45.3 (2010), pp. 641–662.

Table 1: An Open Source Dataset for Asset Pricing

This table compares our dataset to McLean and Pontiff (2016) (MP); Green, Hand, and Zhang (2017) (GHZ); Harvey, Liu, and Zhu (2016) (HLZ); and Hou, Xue, and Zhang (2017) (HXZ). Clear predictors are those where the original papers clearly demonstrate that our portfolios should be statistically significant. Likely predictors were shown to attain significance, but our portfolios are not close enough to the original tests to be confident of a significant reproduction. Maybe predictors had a lack of predictability evidence in the original papers. Not predictors were shown to be statistically insignificant. Widely-available data includes CRSP-Compustat, IBES, OptionMetrics, 13F, FRED, among others. A detailed list of characteristic definitions is in the Online Appendix. Our dataset offers comprehensive coverage of firm-level predictors. Code and data are available at https://github.com/OpenSourceAP/CrossSection.

Panel A: Variable Counts						
	Our Data		Other Metastudies			
	Benchmark	Extended	MP	GHZ	HLZ	HXZ
Firm-Level Characteristics from Widely-Available Data						
Original Characteristics						
Clear predictor	180	180	77	77	227	136
Likely predictor	30	30	12	14	3	19
Maybe predictor		40	8	6	45	33
Not predictor		10		5	4	6
Characteristic Variants		43				46
Additional Portfolios Made From Alternative Rebalancing Frequencies						
		945				212
Other Variables						
Theory					22	
Not Firm-Level Predictors					84	
Less-Available Data					50	
Total	206	1260	97	102	434	452

Panel B: Our Coverage of Other Metastudies (%)				
	MP	GHZ	HLZ	HXZ
Firm-Level Characteristics from Widely-Available Data				
Original Characteristics				
Clear Predictor	98	90	90	100
Likely Predictor	100	100	33	100
Maybe Predictor	100	100	5	100
Not Predictor	100	100	25	100
Characteristic Variants				100

Table 2: Performance of Individual Clear Predictors. This table lists the clear predictors in our baseline data, as well the in-sample mean returns and t-statistics of their corresponding portfolios. Detailed descriptions of predictors are in the Online Appendix, and code and data are at https://github.com/OpenSourceAP/CrossSection. The table is sorted by author and serves as a quick reference guide to our dataset.

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Abarbanell and Bushee	1998	Change in capital inv (ind adj)	1974	1988	0.46	5.29
Abarbanell and Bushee	1998	Gross Margin growth over sales growth	1974	1988	0.35	3.22
Adrian, Etula and Muir	2014	Broker-Dealer Leverage Beta	1973	2009	0.48	3.35
Ali, Hwang, and Trombley	2003	Idiosyncratic risk (AHT)	1976	1997	1.32	3.68
Alwathainani	2009	Earnings growth for consistent growers	1971	2002	0.24	2.60
Amihud	2002	Amihud's illiquidity	1964	1997	0.48	2.73
Anderson and Garcia-Feijoo	2006	Change in capex (two years)	1976	1999	0.48	4.78
Anderson and Garcia-Feijoo	2006	Investment growth (1 year)	1964	2003	0.19	2.68
Anderson and Garcia-Feijoo	2006	Change in capex (three years)	1976	1999	0.51	4.74
Ang et al.	2006	Systematic volatility	1986	2000	1.14	3.58
Ang et al.	2006	Idiosyncratic risk	1963	2000	1.05	3.39
Ang et al.	2006	Idiosyncratic risk (3 factor)	1963	2000	1.05	3.37
Ang et al.	2006	Idiosyncratic risk (CAPM)	1963	2000	1.04	3.31
Asquith Pathak and Ritter	2005	Inst own among high short interest	1980	2002	1.64	2.24
Avramov et al	2007	Junk Stock Momentum	1985	2003	1.22	3.58
Balakrishnan, Bartov and Faurel	2010	Return on assets	1976	2005	0.93	3.19
Balakrishnan, Bartov and Faurel	2010	Return on assets incl extraordinary income	1976	2005	1.38	5.89
Bali, Cakici, and Whitelaw	2010	Maximum return over month	1962	2005	0.81	2.49
Bali, Engle and Murray	2015	Skewness of daily returns	1963	2012	0.49	6.40
Bali, Engle and Murray	2015	Skewness of daily idiosyncratic returns (3F model)	1963	2012	0.34	5.56
Ball et al.	2016	Cash-based operating profitability	1963	2014	0.60	4.35
Banz	1981	Size	1926	1975	1.23	3.59
Barber et al.	2002	Consensus Recommendation	1994	1997	1.61	4.63
Barber et al.	2002	Down forecast EPS	1985	1997	1.02	11.07

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Barber et al.	2002	Up Forecast	1985	1997	0.63	7.00
Barth and Hutton	2004	Change in Forecast and Accrual	1981	1996	0.51	5.33
Bartov and Kim	2004	Book-to-market and accruals	1980	1998	1.34	5.04
Basu	1977	Earnings-to-Price Ratio	1957	1971	0.59	3.69
Bazdresch, Belo and Lin	2014	Employment growth	1965	2010	0.46	5.50
Belo and Lin	2012	Inventory Growth	1965	2009	0.31	4.10
Bhandari	1988	Market leverage	1952	1981	0.41	2.55
Blitz, Huij and Martens	2011	11 month residual momentum	1930	2009	0.74	7.80
Blitz, Huij and Martens	2011	6 month residual momentum	1930	2009	0.38	3.85
Blume and Husic	1972	Price	1932	1971	1.49	3.22
Boudoukh et al.	2007	Net Payout Yield	1984	2003	0.78	2.20
Boudoukh et al.	2007	Payout Yield	1984	2003	0.32	2.34
Bradshaw, Richardson and Sloan	2006	Net debt financing	1971	2000	0.58	8.44
Bradshaw, Richardson and Sloan	2006	Net equity financing	1971	2000	0.68	4.19
Bradshaw, Richardson and Sloan	2006	Net external financing	1971	2000	1.00	5.93
Brennan, Chordia and Subrahmanyam	1998	Past trading volume	1966	1995	0.83	3.02
Chan and Ko	2006	Momentum and LT Reversal	1965	2001	1.39	4.67
Chan, Jegadeesh and Lakonishok	1996	Earnings forecast revisions	1977	1992	0.95	6.98
Chan, Jegadeesh and Lakonishok	1996	Earnings announcement return	1977	1992	1.21	12.98
Chan, Lakonishok and Sougiannis	2001	Advertising Expense	1975	1996	0.70	3.48
Chan, Lakonishok and Sougiannis	2001	R&D over market cap	1975	1995	0.95	5.99
Chandrashekar and Rao	2009	Cash Productivity	1963	2003	0.58	4.01
Chen, Hong and Stein	2002	Breadth of ownership	1979	1998	0.58	4.20
Chordia, Subrahmanyam and Anshuman	2001	Share turnover volatility	1966	1995	0.83	3.81
Chordia, Subrahmanyam and Anshuman	2001	Volume Variance	1966	1995	0.48	3.55
Cohen and Frazzini	2008	Customer momentum	1980	2004	1.12	4.48
Cohen and Lou	2012	Conglomerate return	1977	2009	1.23	6.02
Cohen, Diether and Malloy	2013	R&D ability	1980	2009	0.14	0.73

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Cooper, Gulen and Schill	2008	Asset Growth	1968	2003	1.06	7.80
Cremers and Nair	2005	Governance among blockheld 1	1990	2001	0.35	1.75
Da and Warachka	2011	Long vs short-term earnings expectations	1983	2006	0.57	4.30
Daniel and Titman	2006	Composite equity issuance	1968	2003	0.36	2.73
Daniel and Titman	2006	Intangible return using BM	1968	2003	0.42	2.80
Daniel and Titman	2006	Intangible return using CFtoP	1968	2003	0.42	3.08
Daniel and Titman	2006	Intangible return using EP	1968	2003	0.35	2.52
Daniel and Titman	2006	Intangible return using Sale2P	1968	2003	0.50	2.52
Daniel and Titman	2006	Share issuance (5 year)	1968	2003	0.48	4.11
Datar, Naik and Radcliffe	1998	Share Volume	1962	1991	0.50	3.06
De Bondt and Thaler	1985	Momentum-Reversal	1933	1980	0.59	3.04
De Bondt and Thaler	1985	Long-run reversal	1929	1982	0.82	3.26
Dechow et al.	2001	Short Interest	1976	1993	0.42	3.08
Dechow, Sloan and Soliman	2004	Equity Duration	1966	1999	0.66	4.73
Desai, Rajgopal and Venkatachalam	2004	Operating Cash flows to price	1973	1997	0.33	2.21
Dharan and Ikenberry	1995	Exchange Switch	1962	1990	0.47	3.00
Dichev	1998	O Score	1981	1995	0.63	4.18
Dichev	1998	Altman Z-Score	1981	1995	0.51	3.00
Dichev and Piotroski	2001	Credit Rating Downgrade	1986	1998	0.49	2.10
Diether, Malloy and Scherbina	2002	EPS Forecast Dispersion	1976	2000	0.62	3.07
Doyle, Lundholm and Soliman	2003	Excluded Expenses	1988	1999	0.45	4.38
Easley, Hvidkjaer and O'Hara	2002	Probability of Informed Trading	1984	1998	1.18	4.15
Eberhart, Maxwell and Siddique	2004	Unexpected R&D increase	1974	2001	0.19	2.27
Eisfeldt and Papanikolaou	2013	Organizational Capital	1970	2008	0.47	2.34
Eisfeldt and Papanikolaou	2013	Organizational Capital industry adj	1970	2008	0.62	4.83
Fama and French	1992	Total assets to market	1963	1990	0.64	3.69
Fama and French	1992	Book leverage (annual)	1963	1990	0.27	3.17
Foster, Olsen and Shevlin	1984	Earnings Surprise	1974	1981	0.95	5.24

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Frankel and Lee	1998	Predicted Analyst forecast error	1979	1993	0.50	3.12
Franzoni and Marin	2006	Pension Funding Status	1980	2002	0.35	3.75
George and Hwang	2004	52 week high	1963	2001	0.88	3.62
Gompers, Ishii and Metrick	2003	Governance Index	1990	1999	0.52	2.09
Gou, Lev and Shi	2006	IPO and no R&D spending	1980	1995	0.76	2.26
Grinblatt and Moskowitz	1999	Industry Momentum	1963	1995	0.70	5.48
Hafzalla, Lundholm and Van Winkle	2011	Percent Abnormal Accruals	1989	2008	0.28	4.79
Hafzalla, Lundholm and Van Winkle	2011	Percent Operating Accruals	1989	2008	0.52	4.48
Hafzalla, Lundholm and Van Winkle	2011	Percent Total Accruals	1989	2008	0.39	3.41
Hahn and Lee	2009	Tangibility	1973	2001	0.57	3.66
Hartzmark and Salomon	2013	Dividends	1927	2011	0.40	6.06
Harvey and Siddique	2000	Coskewness	1963	1993	0.03	0.31
Heston and Sadka	2008	Return seasonality	1965	2002	1.02	8.35
Heston and Sadka	2008	Return seasonality years 11 to 15	1965	2002	0.57	6.67
Heston and Sadka	2008	Return seasonality years 16 to 20	1965	2002	0.46	5.07
Heston and Sadka	2008	Returns in not-same month years 16 to 20	1965	2002	0.26	2.35
Heston and Sadka	2008	Return seasonality last year	1965	2002	1.03	8.14
Heston and Sadka	2008	Returns in not-same month last year	1965	2002	0.80	3.26
Heston and Sadka	2008	Returns in not-same years 2 to 5	1965	2002	0.83	3.45
Heston and Sadka	2008	Return seasonality years 6 to 10	1965	2002	0.66	6.63
Heston and Sadka	2008	Returns in different months years 6 to 10	1965	2002	0.57	5.03
Hirschleifer, Hsu and Li	2013	Citations to RD expenses	1982	2008	1.10	3.92
Hirschleifer, Hsu and Li	2013	Patents to RD expenses	1982	2008	1.28	3.71
Hirshleifer et al.	2004	Net Operating Assets	1964	2002	0.75	7.71
Hirshleifer, Hou, Teoh, Zhang	2004	change in net operating assets	1964	2002	1.10	9.95
Hou	2007	Earnings surprise of big firms	1972	2001	0.62	3.94
Hou	2007	Industry return of big firms	1972	2001	2.21	9.52
Hou and Moskowitz	2005	Price delay r square	1964	2001	0.35	2.03

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Hou and Robinson	2006	Industry concentration (Herfindahl) sales	1963	2001	0.21	2.33
Hou, Xue and Zhang	2018	Change in Return on assets	1973	2016	0.32	3.74
Hou, Xue and Zhang	2018	Change in Return on equity	1973	2016	0.36	4.21
Ikenberry, Lakonishok and Vermaelen	1995	Share repurchases	1980	1990	0.33	4.13
Jegadeesh	1989	Short term reversal	1934	1987	2.06	12.67
Jegadeesh and Livnat	2006	Revenue Surprise	1987	2003	0.91	7.23
Jegadeesh and Titman	1993	Momentum (12 month)	1964	1989	1.13	4.46
Jegadeesh and Titman	1993	Momentum (6 month)	1964	1989	1.05	5.17
Jegadeesh et al.	2004	Change in recommendation	1994	1998	0.85	5.15
Johnson and So	2012	Option Volume to Stock Volume	1996	2010	0.70	2.07
Kelly and Jiang	2014	Tail risk beta	1963	2010	0.44	3.14
La Porta	1996	Long-term EPS forecast	1983	1990	0.75	1.98
Lakonishok, Shleifer and Vishny	1994	Cash flow to market	1968	1990	0.80	4.77
Lakonishok, Shleifer and Vishny	1994	Revenue Growth Rank	1968	1990	0.53	3.85
Lamont, Polk and Saa-Requejo	2001	Kaplan Zingales index	1968	1997	0.60	7.03
Landsman et al.	2011	Real dirty surplus	1976	2003	0.48	2.17
Lee and Swaminathan	2000	Momentum and Volume	1965	1995	0.52	2.09
Lev and Nissim	2004	Taxable income to income	1973	2000	0.46	3.27
Li	2011	R&D capital-to-assets	1980	2007	0.77	2.52
Liu	2006	Days with zero trades, past 1 month	1960	2003	0.89	3.95
Liu	2006	Days with zero trades, past 6 months	1960	2003	0.72	3.24
Liu	2006	Days with zero trades, past 12 months	1960	2003	0.91	4.18
Lockwood and Prombutr	2010	Sustainable Growth	1964	2007	0.58	4.66
Loh and Warachka	2012	Earnings streak indicator	1987	2009	0.61	4.44
Loh and Warachka	2012	Number of consecutive earnings increases	1987	2009	1.46	14.24
Lou	2014	Growth in advertising expenses	1974	2010	0.34	3.55
Loughran and Wellman	2011	Enterprise Multiple	1963	2009	0.73	6.36
Lyandres, Sun and Zhang	2008	Composite debt issuance	1970	2005	0.38	5.47

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Lyandres, Sun and Zhang	2008	change in ppe and inv/assets	1970	2005	1.10	9.19
Menzly and Ozbas	2010	Customers momentum	1986	2005	0.63	2.04
Menzly and Ozbas	2010	Suppliers momentum	1986	2005	0.66	2.07
Michaely, Thaler and Womack	1995	Dividend Omission	1964	1988	0.23	2.18
Mohanram	2005	Mohanram G-score	1978	2001	0.43	2.36
Nagel	2005	Inst Own and BM	1980	2003	0.73	2.60
Nagel	2005	Inst Own and Forecast Dispersion	1980	2003	1.02	3.34
Nagel	2005	Inst Own and Idio Vol	1980	2003	0.51	2.80
Nagel	2005	Inst Own and Turnover	1980	2003	1.44	4.03
Nguyen and Swanson	2009	Efficient frontier index	1980	2003	1.49	6.15
Novy-Marx	2010	Operating Leverage	1963	2008	0.35	2.70
Novy-Marx	2012	Intermediate Momentum	1927	2010	0.41	2.45
Novy-Marx	2013	gross profits / total assets	1963	2010	0.46	4.87
Palazzo	2012	Cash to assets	1972	2009	0.40	1.99
Pastor and Stambaugh	2003	Pastor-Stambaugh liquidity beta	1968	1999	0.25	1.53
Penman, Richardson and Tuna	2007	Leverage component of BM	1963	2001	0.19	2.44
Penman, Richardson and Tuna	2007	Enterprise component of BM	1963	2001	0.53	5.48
Penman, Richardson and Tuna	2007	Net debt to price	1963	2001	0.44	3.23
Piotroski	2000	Piotroski F-score	1976	1996	0.72	3.30
Pontiff and Woodgate	2008	Share issuance (1 year)	1970	2003	0.58	4.90
Rajgopal, Shevlin and Venkatachalam	2003	Order backlog	1981	1999	0.46	3.24
Richardson et al.	2005	Change in current operating assets	1962	2001	0.53	5.95
Richardson et al.	2005	Change in current operating liabilities	1962	2001	0.35	4.45
Richardson et al.	2005	Change in equity to assets	1963	2001	0.45	3.46
Richardson et al.	2005	Change in financial liabilities	1962	2001	0.68	11.88
Richardson et al.	2005	Change in long-term investment	1962	2001	0.14	2.39
Richardson et al.	2005	Change in net financial assets	1962	2001	0.54	9.09
Richardson et al.	2005	Total accruals	1962	2001	0.26	2.88

continued on next page

Table 2: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Ritter	1991	IPO and age	1981	1984	1.39	2.70
Ritter	1991	Initial Public Offerings	1975	1987	0.63	2.23
Rosenberg, Reid, and Lanstein	1985	Book to market using most recent ME	1973	1984	1.31	3.73
Rosenberg, Reid, and Lanstein	1985	Book to market using December ME	1973	1984	1.34	3.88
Scherbina	2008	Decline in Analyst Coverage	1982	2005	0.31	3.91
Sloan	1996	Accruals	1962	1991	0.37	4.19
Soliman	2008	Change in Net Noncurrent Operating Assets	1984	2002	0.36	4.51
Soliman	2008	Change in Net Working Capital	1984	2002	0.15	2.00
Thomas and Zhang	2002	Inventory Growth	1970	1997	0.52	5.70
Thomas and Zhang	2011	Change in Taxes	1977	2006	0.32	3.15
Titman, Wei and Xie	2004	Investment to revenue	1973	1996	0.36	4.20
Tuzel	2010	Real estate holdings	1971	2005	0.27	3.21
Valta	2016	Convertible debt indicator	1985	2012	0.28	3.30
Xie	2001	Abnormal Accruals	1971	1992	0.35	5.02
Yan	2011	Put volatility minus call volatility	1996	2005	0.55	3.28
Zhang	2004	Firm Age - Momentum	1983	2001	2.25	5.37

Table 3: Performance of Individual Likely Predictors. This table lists the likely predictors in our baseline data, as well the in-sample mean returns and t-statistics of their corresponding portfolios. Detailed descriptions of predictors are in the Online Appendix, and code and data are at https://github.com/OpenSourceAP/CrossSection. The table is sorted by author and serves as a quick reference guide to our dataset. Likely predictor reproductions have t-stats that center around 2, as one would expect from a close reading of the original papers.

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Abarbanell and Bushee	1998	Sales growth over inventory growth	1974	1988	0.26	2.83
Abarbanell and Bushee	1998	Sales growth over overhead growth	1974	1988	0.36	2.24
Amihud and Mendelsohn	1986	Bid-ask spread	1961	1980	0.61	1.60
Ball et al.	2016	Cash-based operating profitability	1963	2014	0.27	1.74
Barbee, Mukherji and Raines	1996	Sales-to-price	1979	1991	0.70	2.97
Belo, Lin and Vitorino	2014	Brand capital investment	1975	2010	0.53	2.15
Cremers and Nair	2005	Shareholder activism 2	1990	2001	0.25	0.73
Cusatis, Miles and Woolridge	1993	Spinoffs	1965	1988	0.37	1.98
Elgers, Lo and Pfeiffer	2001	Earnings Forecast to price	1982	1999	0.58	2.92
Fairfield, Whisenant and Yohn	2003	Growth in Long term net operating assets	1964	1993	0.19	2.33
Fama and French	2006	operating profits / book equity	1977	2003	0.66	2.69
Fama and MacBeth	1973	CAPM beta	1929	1968	0.71	1.85
Frankel and Lee	1998	Analyst Value	1975	1993	0.13	1.18
Frankel and Lee	1998	Analyst Optimism	1975	1993	0.63	5.31
Franzoni and Marin	2006	Pension Funding Status	1980	2002	0.25	2.64
Frazzini and Pedersen	2014	Frazzini-Pedersen Beta	1929	2012	0.09	0.35
Hong and Kacperczyk	2009	Sin Stock (selection criteria)	1926	2006	0.18	0.99
Hou and Moskowitz	2005	Price delay coeff	1964	2001	0.22	2.40
Hou and Moskowitz	2005	Price delay SE adjusted	1964	2001	0.03	0.27
Hou and Robinson	2006	Industry concentration (Herfindahl) assets	1963	2001	0.17	1.81
Hou and Robinson	2006	Industry concentration (Herfindahl) book	1963	2001	0.22	2.27
Johnson and So	2012	Option Volume relative to recent average	1996	2010	0.53	1.78
Michaely, Thaler and Womack	1995	Dividend Initiation	1964	1988	0.29	1.96

continued on next page

Table 3: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat
Naranjo, Nimalendran and Ryngaert	1998	Dividend Yield	1963	1994	0.48	2.23
Prakash and Sinha	2012	Deferred Revenue	2002	2007	0.38	1.32
Soliman	2008	Change in Asset Turnover	1984	2002	0.33	4.20
Spiess and Affleck-Graves	1999	Debt Issuance	1975	1989	0.13	2.14
Valta	2016	Secured debt	1985	2012	0.13	1.42
Valta	2016	Secured debt indicator	1985	2012	0.08	0.99
Xing, Zhang and Zhao	2010	Volatility smirk near the money	1996	2005	0.56	2.52

Table 4: Additional Characteristics in Extended Data. This table lists the additional original characteristics (not variants) in our extended data set, as well as the in-sample returns and t-statistics of corresponding portfolios. Maybe predictors had a lack of predictability evidence in the original papers. Not predictors were shown to be statistically insignificant. The table is sorted by author and serves as a quick reference guide to our extended data.

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat	Predictor Category
Abarbanell and Bushee	1998	Effective Tax Rate	1974	1988	0.00	0.01	maybe
Abarbanell and Bushee	1998	Change in sales vs change in receivables	1974	1988	0.04	0.45	maybe
Abarbanell and Bushee	1998	Laborforce efficiency	1974	1988	-0.08	-0.93	maybe
Abarbanell and Bushee	1998	Change in gross margin vs sales	1974	1988	0.28	2.68	maybe
Acharya and Pedersen	2005	Illiquidity-illiquidity beta (beta2i)	1964	1999	0.22	1.51	maybe
Acharya and Pedersen	2005	Illiquidity-market return beta (beta4i)	1964	1999	-0.10	-1.17	maybe
Acharya and Pedersen	2005	Net liquidity beta (betanet,p)	1964	1999	0.24	1.57	maybe
Acharya and Pedersen	2005	Return-market illiquidity beta (beta3i)	1964	1999	0.11	0.59	maybe
Acharya and Pedersen	2005	Return-market return illiquidity beta (beta1i)	1964	1999	-0.01	-0.03	maybe
Anderson, Ghysels, and Juergens	2005	dispersion in long-term analyst forecasts	1991	1997	0.19	0.69	maybe
Barry and Brown	1984	Firm age based on CRSP	1931	1980	-0.13	-1.39	maybe
Belo, Lin and Vitorino	2014	Brand capital to assets	1975	2010	0.30	1.54	maybe
Dimson	1979	Dimson Beta	1955	1974	-0.22	-1.43	maybe
Elgers, Lo and Pfeiffer	2001	Number of analysts	1982	1998	0.30	1.22	maybe
Francis, Lafond, Olsson and Schipper	2004	Earnings persistence	1975	2001	-0.22	-1.67	maybe
Francis, Lafond, Olsson and Schipper	2004	Earnings Predictability	1975	2001	0.54	3.17	maybe
Francis, Lafond, Olsson and Schipper	2004	Earnings Smoothness	1975	2001	0.03	0.17	maybe
Francis, LaFond, Olsson and Schipper	2004	Earnings conservatism	1975	2001	-0.01	-0.09	maybe
Francis, LaFond, Olsson and Schipper	2004	Earnings timeliness	1975	2001	0.00	0.06	maybe
Francis, LaFond, Olsson and Schipper	2004	Value relevance of earnings	1975	2001	0.00	0.01	maybe
Francis, LaFond, Olsson and Schipper	2004	RoA volatility	1975	2001	-0.04	-0.11	maybe
Francis, LaFond, Olsson and Schipper	2005	Accrual Quality	1971	2002	0.22	0.82	maybe
Francis, LaFond, Olsson and Schipper	2005	Accrual Quality in June	1971	2002	0.25	0.95	maybe
Frankel and Lee	1998	Intrinsic or historical value	1975	1993	0.97	5.24	maybe

continued on next page

Table 4: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat	Predictor Category
Haugen and Baker	1996	Capital turnover	1979	1993	0.18	0.89	maybe
Haugen and Baker	1996	net income / book equity	1979	1993	0.40	3.44	maybe
Haugen and Baker	1996	Cash-flow to price variance	1979	1993	0.48	1.86	maybe
Haugen and Baker	1996	Volume to market equity	1979	1993	0.49	1.54	maybe
Haugen and Baker	1996	Volume Trend	1979	1993	0.59	2.83	maybe
Holthausen and Larcker	1992	Depreciation to gross PPE	1978	1988	0.29	1.05	maybe
Holthausen and Larcker	1992	Change in depreciation to gross PPE	1978	1988	0.19	1.82	maybe
Hou and Loh	2016	Bid-ask spread based on TAQ data	1984	2012	0.14	0.43	maybe
Ortiz-Molina and Phillips	2014	Asset liquidity scaled by book assets	1984	2006	0.26	0.99	maybe
Ortiz-Molina and Phillips	2014	Asset liquidity scaled by market value of assets	1984	2006	1.42	7.40	maybe
Ou and Penman	1989	CF to debt	1973	1983	0.15	0.56	maybe
Ou and Penman	1989	Current Ratio	1973	1983	0.18	1.31	maybe
Ou and Penman	1989	Change in Current Ratio	1973	1983	0.16	2.00	maybe
Ou and Penman	1989	Change in quick ratio	1973	1983	0.29	3.15	maybe
Ou and Penman	1989	Change in sales to inventory	1973	1983	0.44	4.50	maybe
Ou and Penman	1989	Quick ratio	1973	1983	0.21	1.38	maybe
Ou and Penman	1989	Sales to cash ratio	1973	1983	0.21	1.23	maybe
Ou and Penman	1989	Sales to inventory	1973	1983	0.03	0.16	maybe
Ou and Penman	1989	Sales to receivables	1973	1983	0.31	1.63	maybe
Soliman	2008	Asset Turnover	1984	2002	0.43	2.37	maybe
Soliman	2008	Change in Noncurrent Operating Assets	1984	2002	0.91	5.83	maybe
Soliman	2008	Change in Noncurrent Operating Liabilities	1984	2002	0.51	3.75	maybe
Soliman	2008	Change in Profit Margin	1984	2002	0.14	1.55	maybe
Soliman	2008	Profit Margin	1984	2002	0.68	2.35	maybe
Soliman	2008	Return on Net Operating Assets	1984	2002	0.10	0.85	maybe
Ang, Chen and Xing	2006	Downside beta	1963	2001	0.02	0.10	not
Brown and Rowe	2007	Return on invested capital	1970	2005	0.04	0.16	not
Callen, Khan and Lu	2013	Accounting component of price delay	1981	2006	0.42	1.70	not

continued on next page

Table 4: (continued)

Author(s)	Year	Predictor	Sample Start	Sample End	Mean Return	t-stat	Predictor Category
Callen, Khan and Lu	2013	Non-accounting component of price delay	1981	2006	0.19	1.30	not
Campbell, Hilscher and Szilagyi	2008	Failure probability	1981	2003	0.59	1.33	not
Chan, Lakonishok and Sougiannis	2001	R&D to sales	1975	1995	0.17	0.77	not
Fama and MacBeth	1973	CAPM beta squred	1929	1968	0.71	1.84	not
Heston and Sadka	2008	Returns in different months years 11 to 15	1965	2002	0.22	2.11	not
Richardson et al.	2005	Change in short-term investment	1962	2001	0.06	0.40	not
Whited and Wu	2006	Whited-Wu index	1975	2001	0.49	1.32	not

Figure 1: Pairwise Rank Correlations Between Firm-Level Predictors. Predictors are signed so that a higher value implies higher expected returns. Correlations are pooled across all firm-months available. Data includes both clear and likely predictors. Panel A shows all pairs of predictors. Panel B shows only pairs that include B/M, momentum, and profitability. Our dataset contains many distinct predictors.

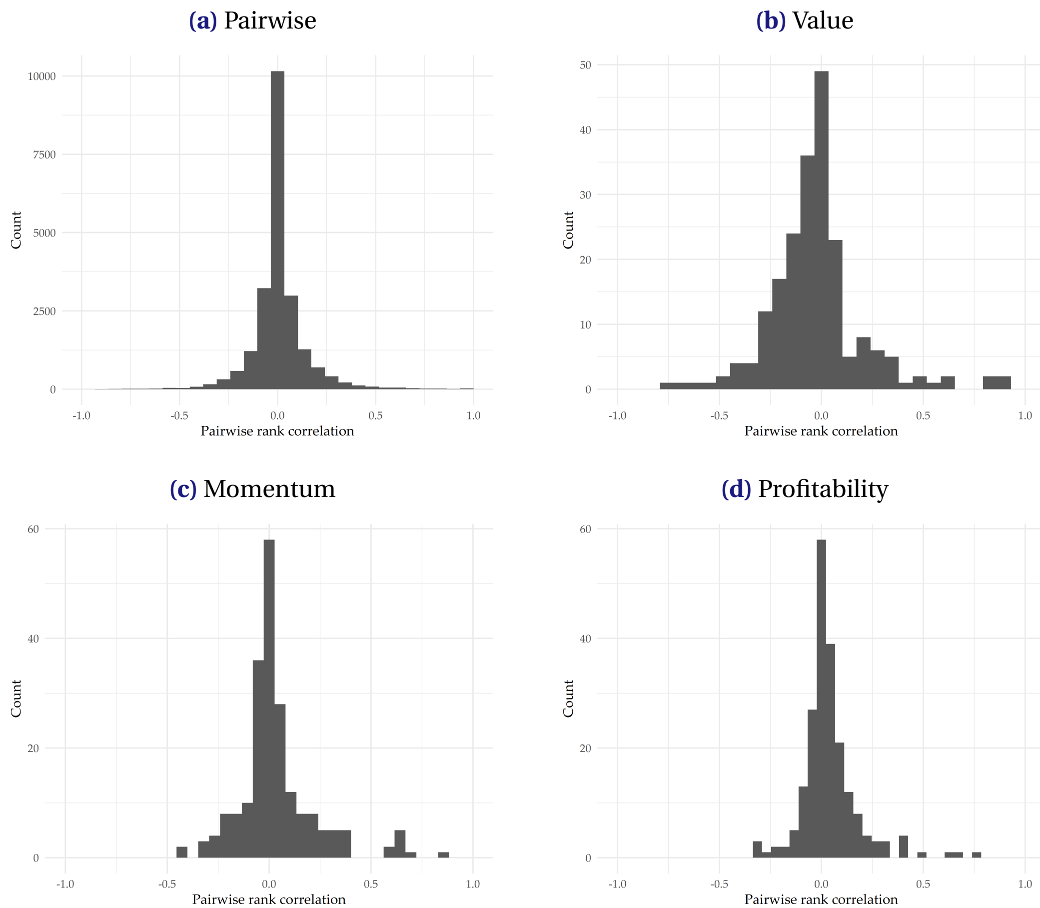

Figure 2: Pairwise Correlations Between Predictor Portfolio Returns. Portfolios are signed to have positive mean returns. Correlations are computed using the longest overlapping samples. Data includes both clear and likely predictors. Our dataset contains many distinct portfolios.

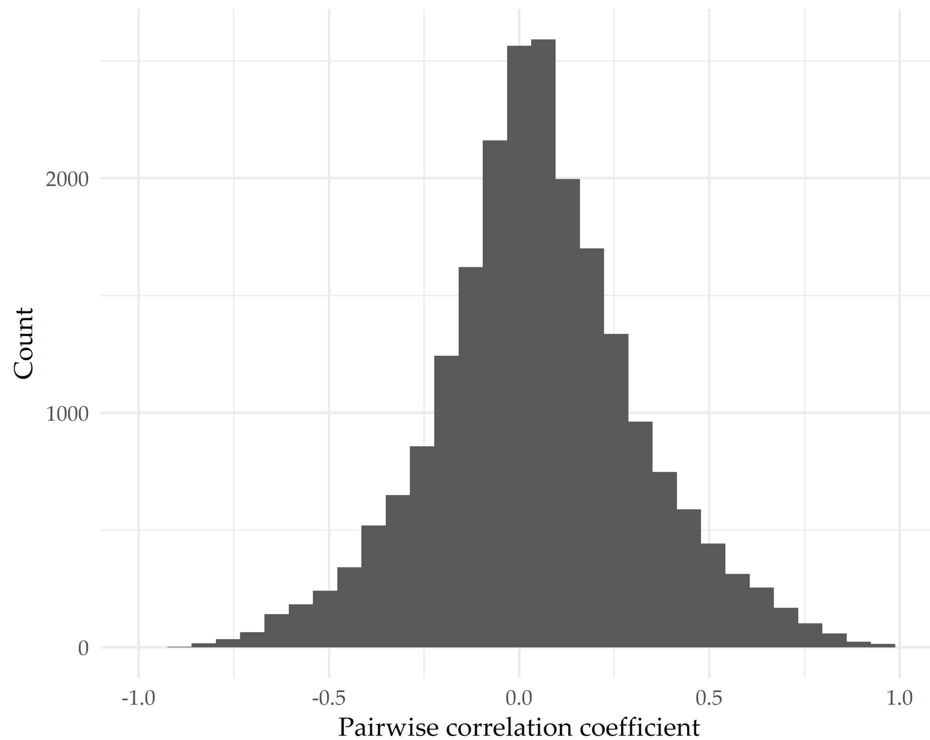

Figure 3: Reproduction Success Rates for Clear Predictors. We construct one long-short portfolio from each clear predictor following the original papers' results and examine the t-stat for the hypothesis that the mean return is zero in the original papers' sample periods. Clear predictors are those where the original papers clearly demonstrate that our portfolios should be statistically significant (see Sections 3.1 and 3.3). Our data reproduces the significance of almost all clear predictors.

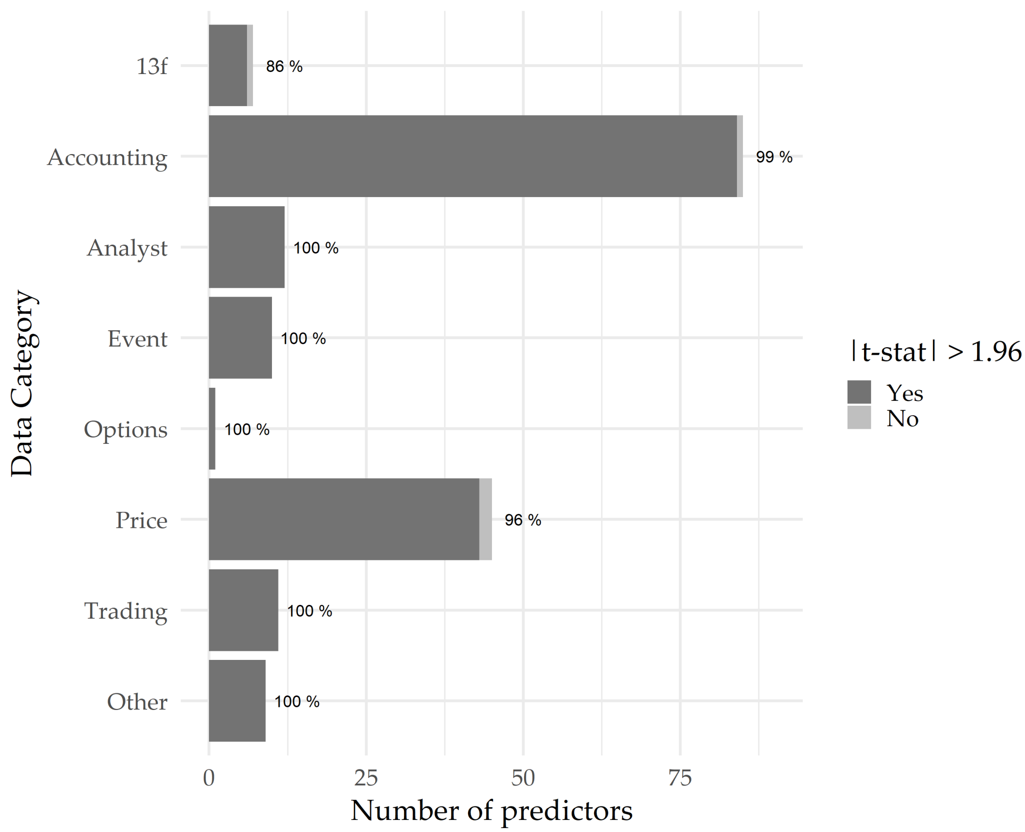

Figure 4: Performance by Rebalancing Frequency. We use baseline characteristics to construct long-short portfolios that take on new signal data every 1-, 3-, 6-, or 12-months and measure mean returns in the original sample periods. We use only characteristics that use 1-month rebalancing in their baseline portfolios. Middle line is median, boxes are 25 and 75 percentiles, and the whiskers extend to the smallest (largest) value within the 25th (75th) percentile minus (plus) 1.5 times the interquartile range. Our code is flexible in rebalancing frequencies and produces the intuitive result that less frequent rebalancing leads to lower mean returns.

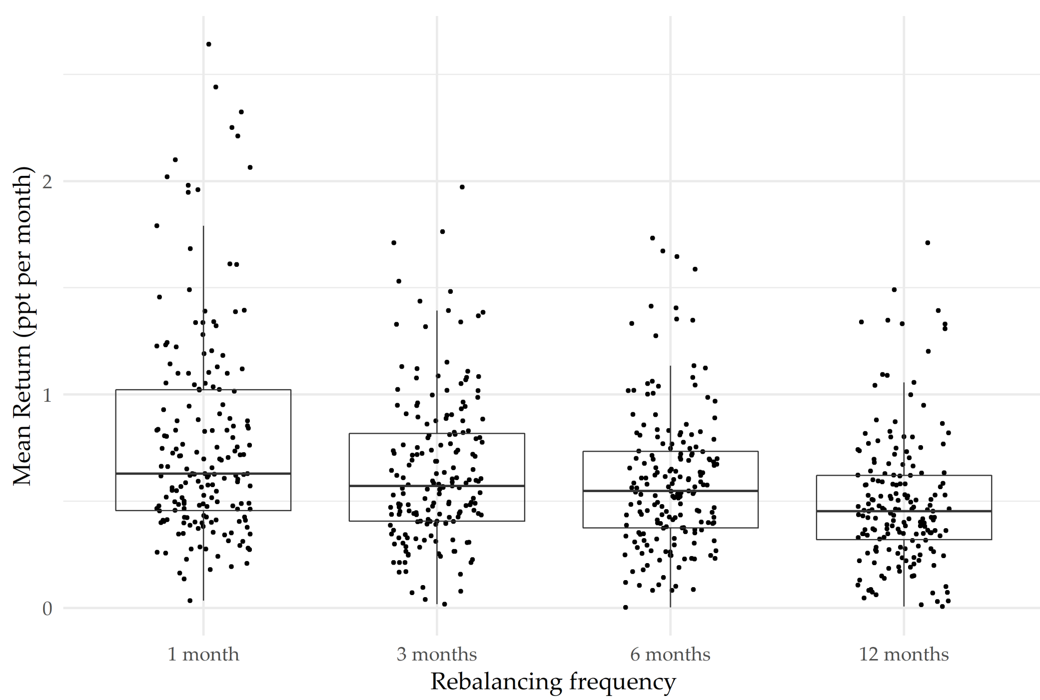

Figure 5: Performance by Liquidity Screen. We use the baseline characteristics to construct long-short portfolios using various liquidity screens: market equity > 20th percentile of among NYSE stocks, NYSE only, and share price > \$5. Each dot is one portfolio. Middle line is median, boxes are 25 and 75 percentiles, and the whiskers extend to the smallest (largest) value within the 25th (75th) percentile minus (plus) 1.5 times the interquartile range. Our code can impose a variety of liquidity screens and produces the intuitive result that liquid stocks are less predictable.

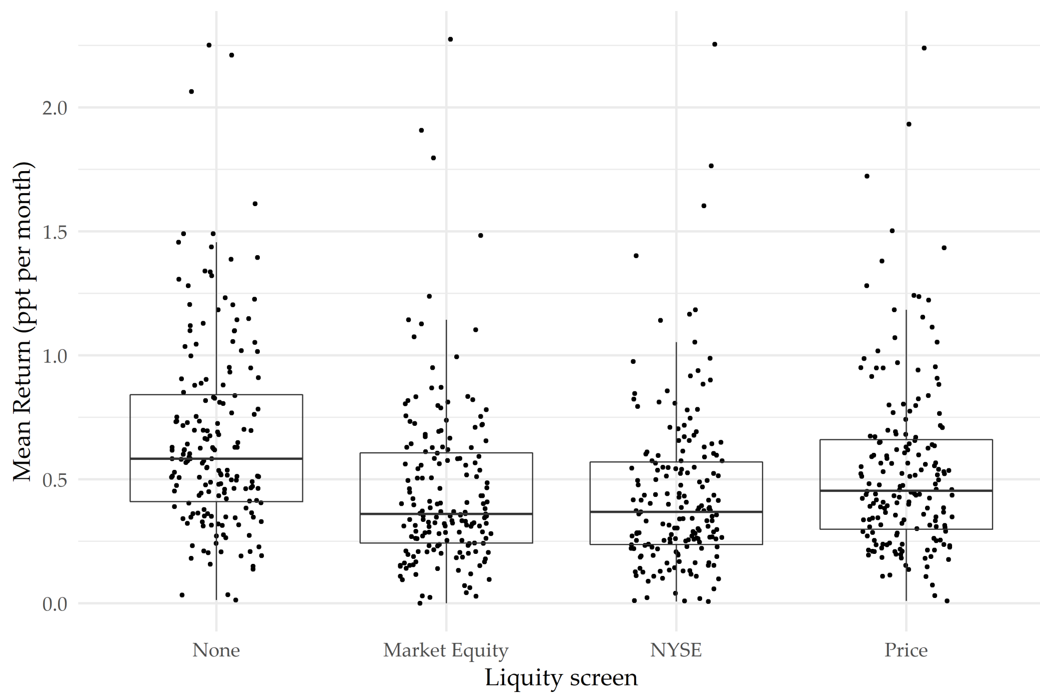

Figure 6: Predictive Performance in Extended Dataset. We construct one long-short portfolio from each characteristic following the original papers and examine the t-stat for the hypothesis that the mean return is zero in the original papers' sample periods. Predictor categories are based on the original papers: clear implies clear evidence our portfolios should be statistically significant, likely predictors had significance, but our portfolios are not close enough to the original tests, maybe predictors had a lack of evidence, and not predictors were shown to be statistically insignificant. Our reproductions find that less-than-clear predictors generally result in t-stats below 1.96.