Fundamental Analysis Works

Söhnke M. Bartram[*] and Mark Grinblatt[†]

Abstract

Stock prices cannot be the outcome of a rational efficient market if fundamental analysis based on public information is profitable. Our approach to fundamental analysis estimates the intrinsic fair values of stocks from the most common quarterly balance sheet and income statement items that were last reported in Compustat. Taking the view of a statistician with little knowledge of the theory of finance, we show that the most basic form of fundamental analysis yields trades with risk-adjusted returns of up to 9% per year. The trading strategy relies on convergence of market prices to their fair values. The greatest rate of convergence occurs in the month after the mispricing signal and subsequently decays to zero over the subsequent 28 months. Profits from trading are present for both large and small firms in economically significant magnitudes.

Keywords: Valuation, asset pricing, market efficiency, fundamental analysis
JEL Classification: G11, G12, G14
This version: March 15, 2015
First version: January 20, 2014

---

[*] Professor of Finance, Warwick Business School, Department of Finance, Coventry CV4 7AL, United Kingdom, Phone: +44 (24) 7657 4168, Fax: +1 (425) 952 10 70, Email: <s.m.bartram@wbs.ac.uk>, Internet: <http://go.warwick.ac.uk/sbartram/>.

[†] Japan Alumni Chair in International Finance, UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles CA 90095-1481, Phone: +1 (310) 573 1635, Email: <mark.grinblatt@anderson.ucla.edu>.

The authors would like to thank especially Juhani Linnainmaa as well as David Aboody, Greg Brown, Adrian Buss, Ian Cooper, Tarun Chordia, Jean Dermine, Bernard Dumas, Peter Feldhütter, Richard Frankel, Federico Gavazzoni, Robin Greenwood, Denis Gromb, David Hirshleifer, Samuli Knupfer, Ralph Koijen, Sonia Konstantinidi, Alok Kumar, John Kuong, Wei-Hsien Li, Karl Lins, Marc Martos-Vila, Anna Pavlova, Neil Pearson, Jeffrey Pontiff, Scott Richardson, Siew Hong Teoh, Raman Uppal, Mathijs van Dijk, Theo Vermaelen, Russ Wermers and seminar participants at Cardiff Business School, Frankfurt School of Finance and Management, INSEAD, London Business School, National Central University Taiwan, Queen Mary University London, University of Illinois at Chicago, the 2014 GEA conference in Kiel, the PhD Nordic Finance Workshop in Stockholm and the 7th Conference on Professional Asset Management in Rotterdam for helpful comments and suggestions. Yuan Shao (Ken) Chang provided excellent research assistance. Kevin Aretz kindly provided data on default probabilities.

One of the cornerstones of market efficiency is the principle that fundamental analysis should "not work." Trading strategies derived from public information like accounting statements should not earn abnormal profits for the risk they bear. Over the past 35 years, evidence has accumulated about anomalies that seem to violate this maxim. Investments linked to momentum, earnings surprises, stock issuance, accruals, credit risk, gross profit, book-to-market, and a host of other signals have earned abnormal profits in the past.[1] However, unlike basic fundamental analysis, the motivation for studying these signals is not always apparent.[2]

Fundamental analysis is based on the principle that stocks have an intrinsic fair value and that investors can earn abnormal profits from stock-specific signals that indicate deviations from fair value. Abnormal profits arise from convergence to fair value – at one extreme via short-term term price movements towards fair value, or more slowly, via distributions of dividends, takeovers, private buyouts, or asset liquidation. Alternatively, to profit from fundamental analysis, one merely has to subscribe to the seemingly plausible hypothesis that share prices are more likely to converge to fair value than diverge from it.

Despite the popularity of the discounted cash flow technique, fundamental analysis does not necessarily require explicit cash flow forecasts and discount rates. These forecasts and discount rates can be implicit in a variety of other approaches that obviate the need for explicit

[1] See, for example, Ball and Brown (1968), Jones and Litzenberger (1970), Joy, Litzenberger, and McEnally (1977), Rendleman, Jones, and Latané (1982), Foster, Olsen, and Shevlin (1984), Bernard and Thomas (1989, 1990), Fama and French (1992), Jegadeesh and Titman (1993), Michaely, Thaler, and Womack (1995), Ikenberry, Lakonishok, and Vermaelen (1995), Sloan (1996), Ball and Bartov (1996), Dichev (1998), Fama and French (2006), Pontiff and Woodgate (2008), Campbell, Hilscher, and Szilagyi (2008), Avramov, Chordia, Jostova, and Phillipov (2009), and Novy-Marx (2013).

[2] Both behavioral and risk-based hypotheses have been advanced to explain anomalies like these, but the explanations have generally been developed after the fact. For example, overconfidence and the disposition effect are offered as behavioral explanations for momentum; return covariation within the value and growth categories, embodied in the HML factor, is proposed as a risk-based explanation for the value premium. See Fama and French (1993), Daniel, Hirshleifer, and Subrahmanyam (1998), and Grinblatt and Han (2005).

models and parameter estimates. We take a particularly simple and agnostic view of how to compute fair value: Rather than select a specific theoretical model, we approximate a stock's fair value as a linear function of virtually all of its most recently reported income statement and balance sheet items. Our only restriction is that the function's coefficients, which are determined each month, offer the lowest degree of mispricing (as measured by variance) of a randomly selected investment dollar in the economy. This more direct approach to fair value estimation is consistent with the most basic principles of asset pricing theory and turns out to be exceedingly simple to implement: fair values are the predictions of monthly cross-sectional regressions of market capitalizations on firm-level accounting data.

This approach to fundamental analysis is unorthodox, but it avoids the temptation to data snoop across model specifications.[3] With the traditional implementations of fundamental analysis, there are not only multiple models of fair value, but countless approaches to earnings forecasts and discount rate estimation. The freedom to define fair value in so many ways leads to a staggering number of investment strategies one could investigate to test the efficacy of fundamental analysis. We believe that the theorist's license is best suspended when it offers too much discretion for selecting the implementation that yields significant results by chance. In contrast, the least squares criterion of the cross-sectional regressions guarantees that the market portfolio is fairly valued at all times, but prevents discretion in the selection or weighting of accounting items that could conceivably relate to future returns.

After identifying fair values from linear functions of accounting items, we study the profitability of buying undervalued and selling overvalued securities to assess whether fundamental

[3] See, for example, Schulmeister (2009).

analysis works. Here, we quantify mispricing as the percentage difference of a stock's actual market capitalization from our estimate of its fair value. With extensive controls for risk and major known anomalies, convergence to fair value is the most likely source of the remarkable profitability we uncover from this trading strategy. The abnormal return (alpha) spreads earned from long-short strategies based on quintile sorts for percentage misvaluation are between 5% and 9% per year, depending on the risk adjustment procedure used, and positive in about 60% of the 432 months studied. They are prevalent in large and small firms, evident in all sub-periods, and not explained by the "usual suspects:" industry returns, beta, book-to-market ratios, momentum, short- or long-term reversals, firm size, gross profitability, accruals, earnings surprises, default risk, and net stock issuance. We use both Fama-MacBeth regressions and Black-Jensen-Scholes time-series factor model regressions to implement these controls.

Our approach to fair value estimation, conveniently referred to as the "statistician's approach" to fundamental analysis, is deliberately crude and made even cruder by the accounting inputs used. The cross-sectional regression essentially uses all balance sheet and income items reported by sufficient numbers of firms. The large numbers of highly (or perfectly) collinear variables implies that coefficient signs will flip month-to-month and many of the variables lack any unique coefficient because they are redundant. More precise ways of obtaining fair values certainly exist, but our goal is to be conservative at assessing whether a crude form of fundamental analysis works. The fair value approach used here is unlikely to be a superior mousetrap for capturing the intrinsic values of securities. If the crude statistician's approach to fundamental analysis works, then more accurate ways of measuring mispricing should work even better.

Despite the handicaps imposed on fair value estimation, our approach has theoretical roots in the most intuitive of principles that guide fair value: the law of one price. Like fair values obtained from any asset pricing model, the fair values obtained with our approach are the market values of replicating portfolios[4] – "replicating" because each of the latter portfolios' accounting items are identical to those of the firm being valued. Because the number of firms N is large relative to the rank K of an NxK matrix X of all firms' accounting data at a given date, an infinite number of portfolios replicate the accounting data of the firm being valued. Each has a distinct market price that represents an estimate of the target firm's fair value. However, as Appendix A proves, among all these fair value candidates, our unique fair value prediction and the replicating portfolio attached to it can be deduced from three appealing assumptions:

1) The portfolio has weights on stocks that make the average valuation error zero (which is equivalent to assuming that the market portfolio is fairly priced).
2) The portfolio weights are functions only of the K-dimensional accounting information and are <u>not</u> functions of firms' market capitalizations, returns, or other variables besides the accounting information.
3) The portfolio minimizes the average squared deviation across securities of <u>any</u> attribute (including market capitalization) not spanned by the K-dimensional accounting attributes.

The set of replicating portfolios satisfying the above criteria forms an NxN idempotent projection matrix $X(X^TX)^{-1}X^T$, tied to the cross-sectional regression described above.

Note that the idempotent projection matrix, and hence the weights of the replicating portfolios, are constructed without regard for any firm's market capitalization! The accounting variable regressors would generate the same idempotent matrix of replicating portfolio weights if it

[4] As Ross (1978, p. 455) pointed out in "A Simple Approach to the Valuation of Risky Streams," even the simplest discounting of risk-free cash flows is merely a comparison between the traded price of a quantity of risk-free bonds available in the securities markets and an asset that produces a future risk-free cash flow. In the CAPM, a stocks fairly valued replicating portfolio is a scaling of the market portfolio and risk free asset with the same beta as the stock. In continuous-time asset pricing, fairly priced Arrow-Debreu securities, constructed from dynamic portfolios of fairly priced assets generate the pricing kernels used to obtain fair values of all assets. And even when parameters like risk aversion are estimated from experiments, the lotteries used to obtain those parameters are deemed to be fairly valued.

was designed to fit earnings growth rates, age of the CEO, or the latitude of the firm's headquarters instead of market capitalization. Despite market capitalization's nonexistent role in the replicating portfolio, the market values of the replicating portfolios capture all of the dynamics of the relationship between market value and accounting variables. And, in contrast to prior studies that predict returns from specific variables of interest, like Price-to-Earnings, Dividend Yield, or Market-to-Book ratios, our valuation approach has little discretion attached to its variables of interest. We are interested in all accounting variables, and any discretion we demonstrate to estimate fair values is based purely on standard statistical criteria – especially, data availability.

As one quantification of the relative importance of our mispricing signal as a return predictor, consider the paper's Fama-MacBeth cross-sectional regressions of returns on the mispricing signal, beta, size, value, and three non-overlapping intervals of past returns representing the past month (short-term reversal), past year (momentum), and past five years (long-term reversal). The mispricing regressor has a coefficient of the correct sign. Its test statistic is of similar significance as that of the momentum regressor, and it surpasses the greater significance hurdles suggested by Harvey, Liu, and Zhu (2013) and by Green, Hand, and Zhang (2013)[5].

Admittedly, higher discount rates imply low market values and vice versa, other things equal.[6] The mechanical relationship in the cross-section between market values (or ratios involving values like book-to-market) and expected returns applies to our signal, as it does to many others in the anomalies literature. However, the mere existence of such a mechanical relationship

---

[5] In Harvey, Liu and Zhu (2013), newly discovered factors should clear a *t*-ratio of 3.00. Green, Hand and Zhang (2013) study more than 330 anomalies and argue that controlling for a subset of existing factors is sufficient for researchers discovering a new predictive factor.

[6] This point was elegantly made by Ball (1978), Berk (1995), and the clean surplus accounting arguments in Fama and French (2006) and Novy-Marx (2013).

does not identify whether the observed return spreads, tied to this type of anomaly, are due to differences in risk or to pricing errors. We present evidence suggesting that convergence to fair value, rather than risk differences, accounts for the efficacy of our mispricing signal.

1 Related Literature

Indirect study of whether fundamental analysis works – measuring the performance of professional money managers – suggests that the abnormal profits earned by those who arguably make a living from fundamental analysis is relatively small: Risk-adjusted returns are in the order of 0-50 basis points per year before deducting transaction costs, fees, and other expenses.[7] Direct study of whether the estimation of fair market values *per se* leads to trading strategies that can earn abnormal profits is rarer. Bhojraj and Lee (2002), Liu, Nissim and Thomas (2002), and Cooper and Lambertides (2014) study the relative valuation of target firms and comparables in the context of multiples valuation. However, Cooper and Lamberion (2014) find no evidence of predictability, and the other two papers do not investigate whether misvaluation can be used to generate profitable trading strategies.

Ou and Penman (1989) study accounting variables as predictors of future earnings changes and show that the probability of an earnings increase predicts stock returns. Abarbanell and Bushee (1998) study the April to March returns of firms with December fiscal year ends and find

[7] See, for example, Grinblatt and Titman (1989, 1993, 1994), Daniel, Grinblatt, Titman and Wermers (1997), Chen, Jegadeesh, and Wermers (2000), Wermers (2000), Fama and French (2010), and Berk and van Binsbergen (2013). Larger performance is achieved when momentum-based returns are not penalized and with international fund management. Jiang, Verbeek, and Wang (2014) show that stocks heavily over-weighted (compared to the index weight) by actively managed funds greatly outperform those heavily under-weighted after adjustment for risk.

that weighted averages of ranks on changes in nine accounting variables predict a firm's return.[8] The discretionary accounting constructs in this latter paper, as well as the weights, are selected because they predict returns in sample. In contrast to our paper, neither of the two aforementioned papers concerns itself with estimation of a firm's fair value and whether deviations from that value have implications for future returns.

The notable exception to the dearth of direct research on fundamental analysis is Frankel and Lee (1998), who study the profitability of trading strategies based on deviations from the residual income model's fair values, obtained from consensus earnings forecasts.[9] In their paper, deviations from fair value predict long-term returns, especially between 24 and 36 month after receiving the mispricing signal. The paper differs from ours in its controls, sample period, and use of specific forecasting model derived from analyst predictions.

The focus of our paper on the relation between a mispricing signal and returns is also markedly distinct from the vast anomalies literature. Research on anomalies has identified more than 300 known predictors of future returns.[10] Their selection, proper mix for a trading signal, and success at publication hinges on the ability of the variable to predict returns. While we relate

---

[8] Related papers include Greig (1992), Holthausen and Larcker (1992), Ou and Penman (1989), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), Piotroski (2000), and Mohanram (2005).

[9] In the same vein, Manaster and Rendleman (1982) show that deviations of observed stock prices from equilibrium stock prices implied in option prices predict future returns for a sample of 172 U.S. stocks. Deviations from fair value have also been used to study misvaluation and Q theories of M&A activity, based on residual income valuation (e.g. Dong, Hirshleifer, Richardson, and Teoh, 2006) or (annual, industry-level) cross-sectional regressions of market capitalization on determinants of fundamental value (e.g. Edmans, Goldstein, and Jiang, 2012; Rhodes-Kropf, Robinson, and Viswanathan, 2005).

[10] See, for example, Green, Hand, and Zhang (2013), Harvey, Liu, and Zhu (2013), and McLean and Pontiff (2013). Kogan and Tian (2013) form pricing factors based on 27 commonly used firm characteristics and show that the relative performance of factor models is highly sensitive to the sample choice and the factor construction methodology, highlighting the challenges of evaluating empirical factor models.

our signal to future returns, the motivation for our hypothesis is uncomplicated and transparent: deviations from fair value are more likely to contract than expand.

2 Data and Methodology for Fair Value Estimation

We now assess whether fundamental analysis from accounting information, implemented with the rudimentary and mechanical approach of a statistician, contains information about future stock returns. At the market close on the last trading day of every month in the sample, we compute each stock's degree of under- or overvaluation. We then track the returns of stocks over the subsequent month[11] and relate these returns to the stock's beginning-of-month mispricing.

2.1 Sample Period and Data Filters

There are 432 return months in our sample: January 1977 through December 2012, and thus 432 portfolio formation dates, starting Friday, December 31, 1976[12] and ending Friday, November 30, 2012. On the day of mispricing measurement and portfolio formation the stock must:

1) Be in CRSP's Monthly Stock File as the only common equity share class of a U.S. corporation (share classes 10 and 11), and be listed on the NYSE, AMEX, or NASDQ-NMS (exchange codes 1-3) with a share price of at least $5 and a positive number of common shares outstanding (to compute market capitalization).
2) Have an earnings announcement date within the past 3 months (for a 10Q or 10K reporting positive total assets) that is at least one trading day prior to the portfolio formation date, according to Compustat.
3) Possess an SIC industry code that is not financial services (SIC codes 60-69).

[11] To compute a return for the month starting at date *t* (also referred to as month *t+1*), we make standard adjustments to the reported CRSP returns for delisting. See, for example, Shumway (1997), Amihud (2002), and Acharya and Pedersen (2005). As delisting is rare, our results are not sensitive to the treatment of delisting.

[12] Quarterly announcement dates are relatively sparse prior to 1976. This dictates the December 31, 1976 start date because portfolio formation requires about one year of these dates.

4) Have two prior fiscal years reported on Compustat.[13]

2.2 Estimating fair value and mispricing

As noted earlier, firm *j*'s date *t* fair value is the prediction, $P_j(t)$, from a cross-sectional regression of firms' actual market values, $V_j(t)$, on accounting variables known by market participants at date *t*. For each of the 432 portfolio formation dates *t*, and each stock *j*, we calculate a mispricing signal, $M_j(t) = [P_j(t)-V_j(t)]/V_j(t)$, the percentage difference between the stock's fair value prediction and its date *t* market capitalization. Underpriced stocks, those with large $M_j(t)$, have low market values relative to the fair values implied by their most recent accounting statements. Such stocks are expected to outperform the overpriced stocks in the future. Conversely, stocks with highly negative $M_j(t)$ are overvalued stocks that are expected to underperform. By construction, the date *t* market-cap-weighted average of $M_j(t)$ is zero.[14]

To economically quantify the effect of mispricing, we rank each of the regression's stocks at the beginning of each month based on the mispricing signal and sort firms into quintile portfolios: Q5 denotes the most underpriced quintile of stock and Q1 the most overpriced quintile. Coefficients from regressing returns on Q2-Q5 dummies can be interpreted as the added return from belonging to the respective mispricing quintile compared to the Q1 quintile.

The regressors for the date *t* fair value predictions come from stock *j*'s (and other firms') most recently reported 10Q or 10K income or balance sheet items, obtained from the CRSP-

[13] This filter, used by Fama and French (1993), Cooper, Gulen and Schill (2008) and Cohen, Polk, and Vuolteenaho (2009), avoids possible IPO-related distortions in the Compustat data.

[14] This property is isomorphic to the fair value regression's average least squares residual beings zero.

merged Compustat Fundamentals Quarterly.[15] We use the prior earnings announcement date from Compustat as the release date of the accounting data. This would be an announcement before December 31, 1976 for the first portfolio formation date in the sample and before November 30, 2012 for the last date in the sample. We deliberately choose the word "before" here because accounting information reported on the last trading day of the month is assigned a disclosure date on the first trading day of the next month. This adjustment controls for month-end information released after the close of trading.[16]

We employ the 28 most commonly reported numerical firm-level[17] Compustat accounting items listed as coming from the balance sheet (14 items) and income statements (14 items) and announced between September 30, 1976 and December 30, 1976. To achieve a 1,000 firm sample at the sample period's start—desirable for statistical precision—28 items is the maximum number we can use.[18] This coverage-imposed reduction of the accounting data matrix, X^*, to 28 columns is fairly innocuous. Many of the uncommon items are redundant—often perfectly or almost perfectly spanned by linear combinations of the more common items. Thus, including additional accounting items adds little to the pertinent valuation information already contained within the most common 28. There is evidence of this even within the 28 items we use. Six of the 28 accounting items are perfectly spanned by the remaining 22. About 98% of the variation

[15] As is customary when analyzing accounting data, all variables that inform trading positions are winsorized – here, based on their a ratio to total assets at the top and bottom 5%, using the sample distribution that exists for that variable from all sample data released prior to month *t*. Our results are not sensitive to winsorization.

[16] To address the possibility that the accounting data we use is not known at Compustat's earnings report date, we rerun our results with earnings report dates artificially pushed forward by 3 trading days. Our results are robust to this change and spot checks of firms' Compustat earnings announcement dates against other sources' dates for the release of accounting data suggest that the occasional differences observed rarely exceed 1 trading day. Using the filing deadlines for 10-Ks and 10-Qs with the SEC as dates when the accounting information was available to investors has also little effect on the results (actual filing dates are only available since 1994).

[17] Many of Compustat's 843 items, like firm name, ticker, and notes are not numerical. Many are titled "per share."

[18] Appendix B lists these 28 items along with details on variables used in the paper as return regression controls.

in half of the items is captured by the remaining half of the items. (While this implies that the regression coefficients on many of the items are imprecise and, in six cases, completely indeterminate, the fair value prediction from the 28 items is unique.)

At the start of each trading month, we use the most recently reported 10K and the three most recently reported 10Qs to identify values for these 28 items. The 14 balance sheet items are from the most recently released accounting statement (10K or 10Q); those for the income statement items are sums of the quarterly values from the three most recently released 10Qs and the most recently released 10K. Although summing four quarterly values characterizes the firm over portions of two fiscal years, it eliminates seasonal distortions that plague the quarterly items themselves. For expositional brevity, the "most recent accounting information" henceforth refers to the 14 items in the most recent balance sheet and the sum of the four quarterly values of the 14 items derived from the four most recent income statements.

Each firm's fair value evolves month to month for two reasons. First, market capitalizations, the cross-sectional regression's dependent variable, change, influencing regression coefficients. For example, rising stock prices imply a larger regression intercept even if accounting information or relative fair values do not change. Changes in relative market capitalizations across market sectors also change these coefficients. When firms with low earnings and large R&D suddenly become more valuable than the historical norm, as in the 1998-99 "internet bubble," our cross-sectional approach will capture that change in market tastes. Second, (in some cases) the firm or other firms may report new accounting information during the month. The new information changes the values and coefficients of the regressors used to predict fair value.

In sum, our approach to fair value takes no stand on changing market preferences for certain types of stocks, or on whether the market as a whole is over- or undervalued at a given point in time. Nor does it rely on a theoretically correct model of fundamental value. Rather, we compare firms to one another. The comparison uses the statistical criterion of goodness of fit to discern how the market values accounting attributes at a given point in time.

All estimates of fair value and mispricing are highly inexact, including ours. The regression's fair market capitalization estimate has R-squareds that vary month-to-month: the minimum R-squared (unadjusted for degrees of freedom) is 74.6% (April 2000), the median is 93.2%, and the average is 91.8%. These R-squareds are unimpressive in light of the fact that market capitalization is on the left hand side and the right-hand side accounting entries tend to scale with firm size. However, there is no bias to the estimation. The noise in our approach only serves to highlight that profits from trading on estimated mispricing can only be improved upon by better mispricing estimation. Nevertheless, the data analysis below will illustrate that even this noisy estimate of mispricing has something interesting to say about market efficiency.[19]

2.3 Summary Statistics for the Overall Sample

Table 1 reports summary statistics describing the relationship of the mispricing variable *M* to firm size, beta, book-to-market ratio, past returns, earnings surprises, accruals, gross profitability, Merton's (1974) default risk, and past returns over a variety of horizons. It reports the time series average of the cross-sectional means of these variables in the first column, the time series average of the correlation of the variable with *M* in the second column, and the times series averages of the means of the variables within five mispricing quintiles. Quintile 1, in the Q1 column, represents the most overpriced stocks, which have equal- and value-weighted average overpricing

[19] Apparently, Numeric Investors, an institutional asset manager, estimates a fair value measure from a similar type of regression on a daily basis, based on cross-sectional regressions of stock price on a proprietary set of company fundamentals (see Perold and Tierney, 1997).

estimates of 223% and 143%, respectively; Quintile 5 (Q5) represents the most underpriced stocks, which have equal- and value-weighted average underpricing estimates of 394% and 286%, respectively.[20]

As can be seen from Table 1, mispricing is highly related to a number of attributes known to predict returns. Compared to the 20% most overpriced firms, the 20% most underpriced firms are about four times smaller, have a lower beta, lower past returns (at all three horizons), and about twice the book-to-market ratio. With respect to size, only about 13 firms among the 20% most underpriced reside in the top size quintile (using NYSE quintile breakpoints), on average. In short, overpriced stocks tend to be long-term winning large growth stocks, with the opposite true for underpriced stocks, but with exceptions. The correlations and quintile averages suggest that there is little relationship between the mispricing metric and the four other firm characteristics we study that are known to predict returns.

The negative correlation between beta and estimated mispricing indicates that beta risk could not explain any ability of M to forecast average returns. Indeed, except for M's positive correlations with value, and negative correlation with firm size and past returns, M seems relatively unrelated to any of the prominent anomalies in the finance literature. We will also control for the effect of book-to-market, size, and past returns on future average returns. The next section shows that M forecasts returns even with controls for these and other effects.

3 The Mispricing Attribute and the Cross-Section of Expected Returns

---

[20] These figures are large because extreme conditional means sort on sampling error and are therefore biased. However, since much of our later analysis involves ranks, there is little need to adjust for the bias with statistical corrections like Bayesian shrinkage.

This paper's assessment of whether fundamental analysis works has two phases. First, as described in the last section, we construct a mispricing signal *M* based on deviation from an estimate of fair value. The second phase, tackled in this section, studies the subsequent monthly returns and abnormal (e.g., risk-adjusted) returns of stocks sorted by the mispricing signal. This second stage determines whether stock prices tend to converge to their fair values over time.

3.1 Raw Returns

Table 2, similar in format to Table 1, addresses the mispricing signal's ability to forecasts next-month's return. It reports time series averages of both equal- and value-weighted portfolio returns in the next month, the average correlation between the return and mispricing signal, as well as the average return of portfolios formed from subgroups of stocks stratified by their mispricing signal. The time series averages are reported both overall and for three similar length sub-periods. In addition to the seven columns from Table 1, Table 2 uses the null of efficient markets to test whether the mean return of the most underpriced quintile of stocks (Q5) exceeds that of the most overpriced quintile. The average difference and associated *t*-statistic, from the time series of paired differences, appear in the two rightmost columns, flanked on their left by the fraction of return differences that are positive.

The average correlation between the signal and future returns is 0.016. Moreover, average returns are also nearly perfectly monotonic in the mispricing quintiles, both for the full sample period and for sub-periods, using both value- and equal-weighted portfolios. The next-month return spread between the least and most underpriced stock quintiles is .68% (.44% when value-weighted), an annualized return spread of 8.1% per year (5.3% per year for the spread in the value-weighted portfolios). Finally, the Q5-Q1 spread is positive in about 61% of the months (56% when portfolios are value-weighted).

3.2 Fama-MacBeth Cross-Sectional Regressions

Table 2's raw return differences could either be due to differences in expected returns associated with the mispricing signal *per se* or to omitted variables linked to the cross-section of returns. To first analyze the issue, Table 3 cross-sectionally regresses firm j's month t+1 return of on the firm's mispricing signal and control variables known at the end of month t. It then averages the coefficients across all months. For a portfolio formed at the end of month t, the cross sectional regression measures the mispricing signal's efficacy from the coefficient $b(t)$ in the regression

$$R_j(t+1) = a_0(t) + b(t)M_j(t) + \Sigma_{\{s=1,S\}}\, c_s(t)X_{js}(t) + error_j(t+1)$$

where

$R_j(t+1)$ = stock j's month $t+1$ return

$X_{js}(t)$= end-of-month t value of firm j's control characteristic s known to influence its returns including industry fixed effects[21]

Table 3's time series averages of coefficients also have t-statistics, outlined in Fama and MacBeth (1973), that appear on their right in parentheses. Table 3 reports several specifications to assess the mispricing signal's ability to predict returns, focusing on three sets of variables. The first is the mispricing signal; the second consists of the more classic characteristics of betas, size, book-to-market and past returns (over three non-overlapping horizons) that correlate with our

[21] On every portfolio formation date, each firm is classified into one of the 38 industries using classifications from the Kenneth French data library, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The regression coefficients and test statistics without industry adjustment or when we force industry fixed effects coefficients to be one negligibly differ from those reported in Table 3.

pricing signal; the third set are the four characteristics of gross profitability, accruals, default risk, and earnings surprises, which mispricing is far less correlated with.[22]

Table 3's first specification lacks controls for other characteristics besides industry, the second and third add the six traditional controls, the fourth and fifth focus on the four additional characteristics known to predict returns, the sixth and seventh control for all other characteristics. Even-numbered specifications exclude the mispricing regressor; odd-numbered specifications include it. Panel A uses the mispricing signal along with the most commonly used functional forms for the controls. Because we do not know the correct functional form for the mispricing signal in this regression, and want to calibrate its economic effect, Panel B reports on analogous regressions using quintile dummies (Q2, Q3, Q4, and Q5, with Q1 omitted due to the regression intercept) for all of the anomalies studied in Panel A. For brevity, Panel B displays coefficients and test statistics only for the Q5 dummy, which represents the difference in returns from being in Q5 compared to Q1; the unreported dummy coefficients for Q2, Q3 and Q4 of each characteristic are also included in the regression.

In all of Panel A and B's specifications, mispricing is a highly significant predictor of next month's return. In Panel A, the smallest *t*-statistic of 3.24 appears in the rightmost "kitchen sink" specification (7). This specification has all of the controls including a number of variables inferred from the firm's accounting statements. In the more traditional Specification 3, our mispricing signal is a more significant predictor of returns (t = 4.42) than logged book-to-market (t = 3.44). Moreover, if one simply transforms the mispricing variable, whose proper functional form

---

[22] To facilitate comparisons across specifications, month *t*'s regressions omit firms lacking data for all specifications. Results are highly similar without this restriction.

is unknown, to a standard normal variable each month, Specification 3's (7's) *t*-statistic becomes 6.12 (3.91).[23]

Table 3 Panel B quantifies the economic effect of the mispricing signal by running the same regression using quintile dummies for all regressors. Here, the mispricing quintile dummy coefficients measure the extra return earned from belonging to mispricing quintile *q* compared to quintile 1 (the overvalued stocks). Panel B's Specification 1 indicates that the average industry-adjusted return of mispricing quintile 5 exceeds that of quintile 1 by 73 basis points per month. *M*'s Q5-Q1 monthly spread drops to 53 basis points with traditional controls (Specification 3), which is two times larger than the book-to-market effect in the same specification, and to 38 basis points in the kitchen sink regression (Specification 7), which is comparable the same specification's momentum effect (37 basis points). In all cases, the mispricing signal is highly significant, and the coefficients are similar if we do not adjust for industry effects.

Clearly, specifications 7's combination of book-to-market with other less traditional controls captures some of the return predictability attributed to mispricing (and momentum). Yet, despite its kitchen sink of controls, our highly constrained approach to fundamental analysis still generates remarkable performance here. The fact that the accounting variables and their weighting are chosen with no input from future returns suggests that a less ad hoc approach to fundamental analysis is likely to prove even more fruitful than what we have proposed.[24]

---

[23] In unreported results, we also find that the mispricing signal predicts the repurchasing and issuing of shares by companies over the subsequent 3-12 months.

[24] Moreover, the dummy coefficients on mispricing signal Q5 average to a positive number in every single calendar month for Specifications 1 and 3, in 10 out of 12 calendar months for Specification 7, and are positive in 60% of the 432-month sample period.

Ruling out one other anomaly helps ensure that our mispricing signal is not "reinventing the wheel." A variety of papers suggest that the net issuance of equity adversely influences future returns.[25] The cross-section of market capitalization used as fair value regression's dependent variable differs from the market capitalization at the time of the accounting statement's fiscal close. This is both because stock prices have changed and because the company has issued or redeemed equity. While we have controlled for past returns, Table 3's Fama-MacBeth regressions have not controlled for net equity issuance. We reran these specifications using fair value regressions that employ market values at the fiscal closing date (on average, 1-2 months prior to the date of market equity employed as the mispricing variable in Tables 1-3). Mispricing estimates based on these more stale market capitalizations yield slightly smaller Q5 dummy mispricing coefficients in Table 3 Panel B, but they remain highly significant. Here, the average Specification 3 coefficient is 0.54 (t=5.85), while Specification 7's coefficient is 0.35 (t=3.72). Moreover, an additional control for issuance in Table 3 results in a similar (sometimes larger) mispricing effect. Thus, omitting a net issuance control does not explain Table 3's mispricing effect.

The last 20 years of our sample period roughly correspond to the two decades in which the profitability of value and momentum strategies became widely known. Panel C of Table 3 averages Panel B's coefficients for the subperiod of 1993-2012. In the four specifications that employ our mispricing signal, the effect of mispricing in the sub-period seems even stronger than in the full sample period. For example, with both the traditional (Specification 3) and kitchen sink controls (Specification 7), the coefficient on the quintile 5 mispricing dummy is about 10 basis points per month larger in Panel C than in B. By contrast, Panel C shows that in the last 20

[25] See, e.g. Ikenberry, Lakonishok and Vermaelen (1995), Loughran and Ritter (1995), Mitchell and Stafford (2000), Teoh and Wong (2002), Schultz (2003), Daniel and Titman (2006), Fama and French (2008), Pontiff and Woodgate (2008).

years, there is no significant value effect in Specification 3 and no significant momentum effect in Specification 7. The mispricing signal is correlated with value and momentum, but it seems to survive a horse race with them. While Specification 7's kitchen sink of controls resurrects a significant value effect, it is largely because gross profitability is a potent predictor of profitable growth stocks' positive returns, while a value strategy predicts that these growth stocks should have poor returns.

3.3 Factor Model Time-Series Regressions

As an alternative to cross-sectional regressions, we estimate factor model alphas of quintile portfolios of firms constructed from the mispricing signal. Compared to cross-sectional regressions, factor models study value-weighted portfolio returns with greater ease and indicate the degree to which long and short positions contribute to the alpha spreads of pairs of quintile portfolios. One disadvantage is that there is wide latitude in how factors are constructed. They may be constructed from averaging across size-stratified groups, are sometimes equal-weighted and sometime value-weighted; even the monthly cross-sectional regression coefficients behind Table 3's results can serve as factors. Factors of a particular design tend to become popular in research when they account for average return anomalies. However, in the seminal formulation of Ross (1976), the selection of factors should be based on their ability to explain covariation, not average returns.

Denote $r_q(t+1)$ to be the industry-adjusted month $t+1$ return on a quintile portfolio based on $Mj(t)$. With L factors, we estimate its alpha as the intercept in the time series regression

$$r_q(t+1) = \alpha_q + \Sigma_{\{i=1,L\}}\beta_{qi}F_i(t+1) + \varepsilon_q(t+1),$$

where $F_i(t+1)$ is the return difference (or excess return) of the i^{th} factor portfolio. If fundamental analysis works, alphas should monotonically increase in the mispricing quintiles. Moreover, the difference in the alphas of the quintile 5 and 1 portfolios – a metric of the mispricing signal's ability to earn abnormal profits – should be significantly positive.

Table 4's industry-adjusted returns are essentially a 0-factor specification. Panel A's industry-adjusted 68 basis point spread between mispricing quintiles 5 and 1 is not identical to the 73 basis point spread in Table 3 Panel B. The spreads differ because Table 4 lifts the requirement that firms possess data for all of Table 3's specifications, and adjusts for industry effects by subtracting the industry return from the dependent variable; Table 3 employs industry dummies as regressors. The industry-adjusted returns are monotonic. Moreover, the 20% most under- and over-priced quintiles exhibit alphas of similar magnitude (but opposite sign) in Panels A and B. Annualized Sharpe ratios here, and throughout the paper, are obtained by multiplying the *t*-statistics of the intercepts by 0.167, the square root of the ratio of 12 to the number of time series observations (typically, 432). For Panel A, the Sharpe ratios of the quintile 5-1 spreads range from 0.98 (0-factor model) to 1.14 (6-factor model). We omit them from the tables for brevity.

Table 4's 6- and 7-factor specifications nest the widely used Fama-French (1993) 3-factor and Carhart (1997) 4-factor models within them. The 6-factor model (Market excess return, SMB, HML, Mom, a short-term reversal factor (ST_Rev), and a long-term reversal factor (LT_Rev)) represents the broadest factor model available in the Kenneth French data library;[26] Table 4's 7-factor model additionally employs the Novy-Marx (2013) profitability factor, PMU,

[26] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The literature does not give unequivocal guidance on the factors that should be included in the risk model. The two additional return factors in the French data library are motivated by the research in DeBondt and Thaler (1985) and Jegadeesh (1990). The results are robust to adding a liquidity factor (from Pastor and Stambaugh, 2003) and a misvaluation factor (from Hirshleifer and Jiang, 2010).

which is now commonly used in investment management.[27] Appendix B provides more detail on all of the factors used in our analysis.

The six factor betas in Table 4's 6-factor model, (all similar to their 7-factor counterparts), indicate that our mispricing strategy is exposed to four of the model's six dimension of factor risk. Compared to overpriced firms, underpriced firms are more exposed to the returns of small value firms with poor medium- and long-term past returns. These findings are similar to those of Table 1 when studying characteristics across the mispricing quintiles.

Table 4's alphas take out the return contribution of these factor exposures. The fairly monotonic alphas of the 6- and 7-factor models in the top and bottom halves of Panels A, respectively, are similar. The monotonicity strengthens the argument that fundamental analysis works. About 70 basis points per month distinguish the two extreme quintile portfolios' alphas, with both the most under- and over-priced quintiles making economically and statistically significant contributions to the alpha spread. This spread, of the same order of magnitude as the raw return spread from Table 2, is also larger than the HML premium, even after controlling for HML!

Panel B weights returns by market capitalization as of the end of month *t*, which is prior to the return month. At 32 basis points per month, the alpha spreads here are more modest. This could indicate that large firms are more fairly priced than our mispricing estimate indicates, warranting firm size as an instrument for shrinking (or stretching) a firm's mispricing estimate with commonly accepted statistical methods. However, the lower spreads for value weighting could also be an artifact of poor diversification. Value-weighted portfolios containing large firms present special inference problems here. When such a portfolio consists of small positions in many

[27] We obtain PMU from Novy-Marx's data library at http://rnm.simon.rochester.edu/data_lib/index.html.

similar-sized firms, firm-specific risk is largely diversified away. The diversified portfolio's return is entirely determined by the portfolio's factor-driven returns plus alpha. Factor model regressors adjust its return for the contribution of factors; moreover, accurate factor beta estimation arises from the portfolio's negligible firm-specific risk. In this case, estimated alphas will be positive, but biased towards zero because fair value estimation imprecisely classifies which firms belong to the extreme mispricing quintiles. However, this bias should be fairly consistent over time if there are many firms to classify as extremely under- or over-priced.

Now consider the same estimation of average returns when a value-weighted portfolio of mispriced firm supplements the many small firms with only a few very large firms. The large firms' firm-specific return component, which cannot be diversified away, contaminates the portfolio's average return and factor beta estimation. This issue is not present in a value-weighted index of all stocks, which achieves diversification by having many large firms. However, with less than 1.5% of its stocks from the top NYSE size decile, it is challenging to estimate the true value-weighted mean return of the most underpriced quintile with any precision. Typically, only one or two of these megacap firms—often 100 times larger than the typical firm—appear in the most underpriced quintile. When this happens, each mega cap's firm-specific risk tends to dominate the return of the value-weighted portfolio and influence the factor betas.

Size-based portfolio sorts using independent sorting procedures – even with equal weighting – do not overcome the inference problem. For example, in the top NYSE size quintile, the alphas of the 20% most underpriced stocks overall exceed those of the 20% most overpriced by 83 and 86 basis points per month when benchmarked against the 6- and 7- factor model, respectively. Yet spreads of this size, despite being rare in our reading of the literature, fail to attain the 5% significance threshold. The insignificance stands in contrast to the significant negative alphas of the overpriced firms in the same size quintile, despite being of smaller magnitude than the 83 and 86 basis point spreads. The paradox is resolved by recognizing how few highly underpriced firms there are within the largest NYSE quintile of stocks. One cannot have a diversified portfolio with only a handful of stocks. By contrast, many stocks within the top NYSE quintile are estimated to be over-priced.

Investigating size-based results using a sequential sort on size and mispricing, which we study next, has the advantage of more firms and thus more diversification within the most underpriced quintile of the largest firms. However, sequential sorts have the disadvantage of measuring alpha spreads for a smaller mispricing difference, particularly among large cap stocks, implying smaller alpha spreads for this category.

3.4 Results for Firms Sorted by Size and Value

To better assess the role of size and value in the efficacy of the mispricing signal, Table 5 Panel A reports industry-adjusted returns and alphas of equal-weighted portfolios of the 20% most under- and overpriced firms within each of 5 size baskets. The sequential sorts of size and mispricing use monthly NYSE quintile breakpoints for market capitalization. Panel B is similar, but for firms sorted first into equal-sized value quintiles and then into mispricing quintiles. In the Q5-Q1 rows of both panels are the average industry-adjusted return or alpha spreads between the 20% most under- and over-priced stocks within each of the size and book-to-market quintiles.

Table 5's industry-adjusted return and alpha spreads between the 20% most under- and over-priced stocks are large for every size category and for the three most value-oriented categories. Panel A's industry-adjusted return spreads range from 41 (largest NYSE quintile) to 77 (smallest NYSE quintile) basis points per month. The comparable ranges for the alpha spreads are 41-90 (6-factor) and 30-90 (7-factor) basis points per month. All three metrics exhibit almost perfect monotonicity in firm size, driven more by the monotonicity of underpriced firms' returns and alphas. All 15 of Panel A's Q5-Q1 *t*-statistics (in parentheses) are highly significant.

The largest NYSE quintile tends to have substantially (and significantly) lower spreads from our mispricing strategy, but only because the sequential sort on size and mispricing generates lesser estimated underpricing in the largest firms' most underpriced quintile of stocks. Table 5 Panel A's mean percentile and quintile rank column offer evidence supporting this assertion. On average, the underpriced firms within the largest size category are at the 62nd underpricing percentile overall. By contrast, in the smallest size category, they are at the 93rd percentile. The average mispricing quintile of the 20% most underpriced stocks in the largest NYSE quintile is a mere 3.62, whereas for the smallest NYSE quintile, the overall quintile rank is 4.95. Thus, the gap in mispricing is narrower for large firms because of the sequential sort.

To test whether size per se influences the mispricing signal's efficacy, we compare "predicted" industry-adjusted returns and alphas with the actual values. The numbers in the "predicted" column use each category's probability distribution of the full sample quintile ranks as weights for the corresponding risk-adjusted returns and alphas in Table 4 Panel A. For example, among May 1997's largest stock quintile, the 20% most underpriced has 0%, 0%, 50%, 41% and 9%, that are in underpricing quintiles 1-5, respectively (compared to the overall sample that month). Table 4 Panel A provides 6-factor alphas for each of the five underpricing quintiles. Weighting those five alphas by 0%, 0%, 50%, 41% and 9% generates a predicted alpha for May 1997. As these percentages change, the prediction changes, generating a time series of predicted alphas, which we ultimately average for Table 5's "predicted" columns.

For the largest size category, the time series average of the predicted industry-adjusted return, 6-factor alpha, and 7-factor alpha of the category's most underpriced quintile of firms do not significantly differ (at the 5% threshold) from the actual values of –1.8, 13.8, and 2.7 basis points per month. Indeed, for the three largest size quintiles, the predicted spreads of all three performance metrics do not differ significantly from the actual spreads. Thus, large size per se does not seem to dampen the efficacy of the mispricing signal seen in Table 4. On the other hand, in the two smallest size categories, all four of the actual alpha spreads significantly exceed the predicted alpha spreads obtained from the distribution of mispricing quintile ranks. Thus, our mispricing strategy may work better for small firms than Table 4's full sample results indicate.

In contrast to the mispricing signal's ability to generate significant Q5-Q1 spreads across all size categories, the Q5-Q1 spreads do not significantly differ from zero for the two most growth-oriented quintiles in Table 5 Panel B. The insignificance stems from the two categories lacking positive risk-adjusted returns. In contrast to the size effect, the lower returns and alphas of these two categories, particularly with the 7-factor model, do not appear to be an artifact of the sequential sort. On average, the sorts' most underpriced firms are in the 65th (lowest book-to-market quintile) and 78th (2nd lowest book-to-market quintile) overall, with average quintile ranks of 3.73 and 4.36, respectively. However, the distribution of ranks (versus the full sample) warrants negligible 7-factor alphas, not the sizable −40 and −23 basis point 7-factor alphas observed.

4 Convergence to Fair Value Better Explains the Results than Alternatives

The term "mispricing signal" implies that it is the subsequent convergence to fair value of mispriced stocks—rather than signal-related risk differences—that accounts for the abnormal returns documented in this paper. This section analyzes whether risk differences are a reasonable alternative explanation for the return spreads observed from our mispricing signal.

4.1 Evidence in the paper already controls for known sources of risk

The lower market betas for underpriced stocks (Table 1) represent the first piece of evidence that our signal works because of mispricing and not risk. With multi-factor risk, one compares Specification 1 in Table 3 Panel B with Table 4 Panel A. Specification 1 indicates that the most underpriced quintile's industry-adjusted returns exceed those of the most overpriced quintile by 73 basis points per month. Table 4's alphas control for known sources of multi-factor risk. However, controlling for these factor exposures has virtually no impact on the industry adjusted return spread: alphas are 71 and 69 basis points per month with the 6- and 7-factor models, respectively.

4.2 The Fama-MacBeth Cross-Sectional Coefficients Have No Known Factor Risk

Table 3 Panel B's signal Q5 dummy coefficients control for other characteristics. Fama and MacBeth (1973), among others, note that the coefficient quantifying this effect is the time series average of self-financing portfolio returns. If Z_t is the matrix (number of firms by regressors) representing data of month t's regressors, the signal 5 dummy coefficient is the product of the portfolio weights given by the corresponding row of $(Z_t^T Z_t) Z_t^T$ and the column vector of the firms' month t returns. This time series of self-financing portfolio returns holds other characteristics fixed, but could correlate with factor portfolio returns formed from these characteristics. When regressing the times series of these portfolio returns (equivalently, the time series of coefficients) on the factors, the intercept (alpha) negligibly differs from the coefficients in Table 3 Panel B. For example, the coefficient of 53 basis points per month for the signal Q5 dummy in Specification 3 of Table 3 Panel B generates alphas of 56, 58 and 54 basis points per month when the time series of coefficients is regressed on Table 4's 6-factor model, Carhart's (1997) 4-factor model, and Fama and French's (1993) 3-factor model, respectively. For Specification 7, Table 3 Panel B's coefficient of 39 basis points per month generates 7-factor, 6-factor, 4-factor, and 3-factor alphas of 40, 38, 39, and 35 basis points, respectively. These similarities arise because the returns implicit in the Fama-MacBeth coefficients have mostly negligible factor betas.[28] Thus, Table 3's controls for other characteristics largely eliminate known factor risks.

4.3 Undiscovered sources of risk

Characteristics that correlate with market values in the cross-section, like book-to-market, size, or long-horizon past returns may explain the cross-section of expected returns. Other things equal, size is inversely related to risk in a rational market that values stocks as the discounted stream of future dividends. After all, discount rates are expected returns. Our mispricing metric is based on the percentage deviation of a crude fair value estimate from market capitalization. The lower market capitalizations of underpriced firms are isomorphic to larger discount rates for future dividend streams, other things equal. This point, from Berk (1995), is a tautology. But there is nothing in the tautology (or Berk's argument) suggesting that high discount rates have to come from higher risk rather than irrational sentiment. Only data analysis and calibrations can distinguish the risk-based and sentiment-based explanations for our findings.

If our estimate of mispricing "works" because it proxies for an omitted risk factor related to market capitalization, then controlling for size-related risk factors should largely eliminate the abnormal returns earned. However, the 6-factor model gives the same abnormal returns as no controls. Controlling for size-based characteristics rather than factor exposures eliminates about 25% of the paper's documented abnormal returns (as evident in Specification 3 of Table 3 Panel B). However, it still leaves an abnormal return that is on the order of the market risk premium.

[28] The only significant factor beta here is on the long-term reversal factor. However, the average factor beta (about 0.1) and the relatively small premium of this factor imply that this loading can only alter alpha to a small degree.

Another perspective on the omitted risk factor hypothesis is offered by signals that are more tied to market capitalization and less to the accounting variables that determine fair value. Table 6 reports industry-adjusted returns along with 6- and 9-factor alphas for a pseudo signal constructed as the percentage deviation of "pseudo fair value" from market capitalization. The pseudo signal drops all 28 accounting variables and thus regresses market capitalization only on a constant.[29] Once we control for SMB and other standard risk factors, we no longer find risk-adjusted returns from the "pseudo-signal." By contrast, controlling for these risk factors, including SMB, has little effect on the mispricing signal's return spreads studied throughout this paper.

Finally, if our results are explained by an omitted risk variable tied to cross-sectional differences in size, stale signals should produce almost the same abnormal returns as fresh signals of estimated mispricing. Cross-sectional differences in book-to-market ratios take years to dissipate. Hence, return differences across firms based on book-to-market ratios are similar irrespective of whether book-to-market ratios are measured at the end of the prior month or one year. By contrast, our accounting-based signal generates ranks that decay more rapidly. The average Spearman rank correlation between the vector of mispricing at month *t* and at *t-1* is 0.92, while the same correlation for the book-to-market ratio is 0.98. Moreover, the rank correlation between months *t* and *t-12* is 0.53 for our mispricing measure, while it is 0.81 for the book-to-market ratio. Hence, if our results were generated by differences in an omitted risk variable, that risk attribute has to change rapidly, and it seems unlikely to be due to a more stable characteristic, like the cross-sectional difference in firm size. Buttressing this argument is evidence on the efficacy of the signal as the signal becomes staler—a topic we turn to next.

[29] For fair comparisons with our other results, we require firms to have non-missing data on the 28 variables even though we do not use them in the fair value regressions.

4.4 Signal Delay

The quarterly data used to construct our mispricing signal is constantly being refreshed. Estimated mispricing ranks change monthly as some firms report new accounting data and market values change. If the abnormal returns we observe from our mispricing estimate arise from stock prices converging to their fair values, stale signals should be less valuable than fresh signals of fair value. To test this hypothesis, we lag the mispricing signal by up to three years. At a two and a half year lag, most profitability from the signal has disappeared. Figure 1 graphs evidence supporting this hypothesis. Using (for fair comparisons) returns beginning in January 1980, Figure 1 graphs the 6- and 7-factor alpha spreads for equally weighted portfolios of stocks in the extreme mispricing quintiles. These stocks are grouped into quintiles based on lags of the mispricing signal ranging from 0-35 months.[30] The decay in the signal's efficacy is rapid in the first three months. For example, with the 6-factor alpha, the signal's initial ability to earn abnormal returns of 70 basis points over the next month drops to 53 basis points when the mispricing signal is a month old. Between the time the signal is 3 and 12 months old, it can generate only 27 basis points a month, and then only 20 basis points per moth the following year. The 7-factor alphas pattern is highly similar.

Figure 1's evidence on the diminished efficacy of delayed mispricing signals is difficult to reconcile with an omitted risk factor as a source of the efficacy. In particular, if this omitted risk factor is correlated with the market capitalizations in the mispricing signal, the signal's efficacy should not decay so rapidly. The cross-section of market capitalization is relatively stable, and the mispricing ranks of stocks largely change because of changes in relative fair values, not

[30] Because of the later start date, the alphas for the 0 lag point differs slightly from the corresponding performance metrics reported in Table 4 Panel A.

because of changes in market capitalizations. The autocorrelation coefficients for the cross-section of stocks' signal ranks are 0.918, 0.855, and 0.801 at lags 1-3, but are 0.965, 0.939, and 0.917 for fair market values, and 0.996, 0.993, and 0.990 for market capitalizations. Hence, for omitted risk to explain why fresh signals are much more profitable than moderately stale signals, it must be a risk tied to the accounting data and not to each firm's market capitalization.

Figure 2 shows the 6- and 7-factor alphas when updating market capitalization and accounting data, but using stale regression coefficients for weighting the accounting variables to derive fair value. Using weights that are one year old reduces performance by about one third. While both the stale (and most recent) coefficients are estimated with error, averaging the weights over various windows from the past does not prevent the drop in performance, lending further support for the conjecture that mispricing is an anomaly rather than a risk factor.

4.5 The Relaxed Investor Who Reduces Turnover

The signal delay results help estimate the profitability of a strategy that places signal-based trades and holds them for a full year. In a steady state, a strategy that puts on positions once, estimated more efficiently with overlapping one year returns, is like an equal-weighted combination of 12 strategies obtained from lags for the signal ranging from 0 to 11. The average alphas from such a "relaxed strategy," as measured by averaging the first 12 alphas in Figure 1—namely, 34 basis points per month for both the 6- and 7-factor models—have far lower turnover than a strategy that holds its signal-induced positions for only one month. With a signal that is refreshed every month, a long-short mispricing strategy in the extreme quintiles has turnover of 226% per year, whereas holding positions for one year leads to annual turnover of 59%. There are negligible differences between the turnover ratios of the long and short positions.

Table 7 more properly derives full-sample abnormal returns and test statistics for the 1-year hold positions of the relaxed investor using Jegadeesh and Titman's (1993, 2001) technique. This approach commences the calculation of returns starting in December 1977 (versus 1979 in the paragraph above) with signals from December 1976 (11 month delay) to November 30, 1977 (0 delay). While 2012 signals that are k months prior to the end of 2012 are missing returns for the last $12 - k$ months of the year-long holding period, this is offset by the 1977 signals that are k months before November 30, 1977 (with returns only for the last $12 - k$ months of the 1-year period). Table 7 suggests that portfolio revisions on a yearly rather than a monthly frequency leads underpriced firms to outperform the overprice firms by 34-35 basis points per month.

4.6 Revisions to Accounting Items

We also investigate whether there is a potential bias in the results due to the fact that Compustat's accounting numbers reflect restated information that was not available to the market at the time. To this end, we re-run all results using the first reported accounting numbers.[31] The performance of our trading strategy using as-first-reported accounting information is only marginally weaker (by 8-9 bp per month) than the performance obtained using restated accounting information, suggesting that the effect of any restatement bias is small (Table 8).

5 Conclusion

Regression-based fitting of accounting data to stock values leads to a mispricing measure constructed from regression residuals. The predicted values are the market values of replicating port-

[31] Note that the comparison between Compustat's As First Reported (AFR) data base, which is rarely used in the research literature, and the regular Compustat file can only commence in 1988 and is based on 24 of the 28 accounting items due data availability.

folios that are assumed to be fairly valued. Ranking firms based on their residual-implied percentage mispricing measure predicts returns in the subsequent month and up to two and a half years in the future. The results are not related to the most commonly known predictors of the cross-section of expect returns. Abnormal return spreads based on mispricing metrics formed from accounting data are in the order of 4-9% per year.

Our approach to fundamental analysis uses only information in the most recent accounting statements to see if prices reflect this information. We find they do not. One can earn risk-adjusted returns of a magnitude earned by value and momentum strategies with rudimentary statistical analysis of the most commonly reported accounting information. One could, of course, investigate other potentially valuable information with the type of statistical analysis undertaken here. The other information could include changes in the same item in consecutive accounting statements, analyst forecasts, or corporate actions. It could even combine this information with the information in past price movements. We leave exploration of other sources of information to future research. Our task here was to examine if the most rudimentary form of fundamental analysis works. It seems to work very well, indeed!

Perhaps the most controversial aspect of our results is the claim that the profits obtained are from fundamental analysis. By using the term "fundamental analysis," we are ultimately telling a behavioral story about mispricing arising from convergence to fair value. We have, however, presented evidence supporting the claim that the abnormal profits earned from fundamental analysis are not due to an omitted risk factor.

We focus only on returns, adjusted for risk factors, rather than more direct measures of convergence, because measuring convergence from returns is a more conservative approach. Our estimate of fair value exhibits regression towards the mean over time, like most other estimates.

Hence, direct measure of the dynamics of the distance between fair value and prices leads to stronger convergence estimates than examining returns alone. Holding fair values constant, underpriced stocks that witness price increases and overpriced stocks that witness price decreases converge to the old fair value. Holding prices fixed, fair value estimates that greatly exceed prices tend to decline while those well below the same price tend to increase. Hence, the simple regression to the mean phenomenon implies that direct measurement of convergence is a less conservative approach for making the point that fundamental analysis works, and it has continued to work for more than 35 years.

Because we focus indiscriminately on the most available accounting items, and because of their high degree of collinearity, which renders the exercise extraordinarily difficult, we have not identified which accounting variables are best for determining fair value. Addressing this question also violates the Occam's razor approach of the naïve statistician used here to obtain fair values. Because the accounting data seems to have a factor structure underlying it, it would not surprise us if only a handful of accounting variables could do as well, or improve upon, the strategies derived here. We leave that, as well as improvements in the fair value estimation approach, to future research.

References

Abarbanell, J. S., and B.J. Bushee, 1997. Fundamental analysis, future earnings, and stock prices, Journal of Accounting Research 35(1), 1-24.

Abarbanell, J.S., and B.J. Bushee, 1998. Abnormal Returns to a Fundamental Analysis Strategy, Accounting Review 73:1, 19-45.

Acharya, V., and L.H. Pedersen, 2005. Asset Pricing with Liquidity Risk, Journal of Financial Economics 77, 375-410.

Altman, E.I., 2012. Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta Models," Handbook of Research in Empirical Finance, edited by C. Brooks, E. Elgar, Cheltenham, UK.

Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-section and Time-series Effects, Journal of Financial Markets 5, 31-56.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov, 2009. Credit Ratings and The Cross-Section of Stock Returns, Journal of Financial Markets 12, 469- 499.

Ball, R., 1978. Anomalous relationships between securities yields and yield surrogates. Journal of Financial Economics 3, 103-126.

Ball, R., and E. Bartov, 1996. How naive is the stock market's use of earnings information?, Journal of Accounting and Economics 21, 319-337.

Ball, R., and P. Brown, 1968. An empirical evaluation of accounting income numbers, Journal of Accounting Research, 159-178.

Berk, J., 1995. A critique of size-related anomalies, Review of Financial Studies 8, 275-286.

Berk, J., and J.H. Van Binsbergen, 2013. Measuring Skill in the Mutual Fund Industry, Stanford University Working Paper.

Bernard, V.L., and J. Thomas. 1989. Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium?, Journal of Accounting Research 27, 1-36.

Bernard, V.L., and J. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, Journal of Accounting and Economics 13, 305-340.

Bhojraj, S., and C.M.C. Lee, 2002. Who is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. Journal of Accounting Research 40:2, 407-439.

Bradshaw, M.T., S.A. Richardson, and R.G. Sloan, 2006. The relation between corporate financing activities, analysts' forecasts and stock returns, Journal of Accounting and Economics 45, 53-85.

Campbell, J.Y., J. Hilscher, and J. Szilagyi, 2008. In Search of Distress Risk, Journal of Finance 63:6, 2899-2939.

Carhart, M., 1997. On Persistence in Mutual Fund Performance, Journal of Finance 52:1, 57-82.

Chen, H. L., N. Jegadeesh, and R. Wermers. 2000. The value of active mutual fund management: An examination of the stockholdings and trades of fund managers, Journal of Financial and Quantitative Analysis 35:3, 343-368.

Cohen, R., C. Polk, and T. Vuolteenaho, 2009. The Price is (Almost) Right, Journal of Finance 64:6, 2739-2782.

Cooper, I.A., and N. Lambertides, 2014. Is there a limit to the accuracy of equity valuation using multiples? London Business School Working Paper.

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

Daniel, K., and S. Titman, 2006. Market reactions to tangible and intangible information, Journal of Finance 61, 1605–1643.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997. Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, Journal of Finance 52, 1035-1058.

Daniel, K.D., D. Hirshleifer, and A. Subrahmanyam, 1998. A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-reactions, Journal of Finance 53, 1839-1885.

DeBondt, W., and R. Thaler, 1985. Does the Stock Market Overreact?, Journal of Finance 40, 793-805.

Dichev, I.D., 1998. Is the Risk of Bankruptcy a Systematic Risk?, Journal of Finance 53,1131-1147.

Dong, M., D. Hirshleifer, and S.H. Teoh, 2012. Overvalued Equity and Financing Decisions. Journal of Accounting and Economics 25:12, 3645-3683.

Dong, M., D. Hirshleifer, S. Richardson, and S.H. Teoh, 2006. Does Investor Misvaluation Drive the Takeover Market. Journal of Finance 61:2, 725-762.

Edmans, A., I. Goldstein, and W. Jiang, 2012. The Real Effects of Financial Markets: The Impact of Prices on Takeovers. Journal of Finance 67:3, 933-971.

Fama, E.F., and J.D. MacBeth, 1973. Risk, Return, and Equilibrium: Empirical Tests, Journal of Political Economy 81: 3, 607-636.

Fama, E.F., and K.R. French, 1992. The Cross-Section of Expected Stock Returns, Journal of Finance 47, 427-465.

Fama, E.F., and K.R. French, 1993. Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics 33, 3-56.

Fama, E.F., and K.R. French, 1996. Multifactor explanations of asset pricing anomalies, Journal of Finance 51, 55-84.

Fama, E.F., and K.R. French, 2006. Profitability, investment, and average returns, Journal of Financial Economics 82, 491-518.

Fama, E.F., and K.R. French, 2008. Average returns, B/M, and share issues, Journal of Finance 63:6, 2971-2995.

Fama, E.F., and K.R. French, 2010. Luck versus skill in the cross-section of mutual fund returns, Journal of Finance 65, 1915–1947.

Ferguson, M.F., and R.L. Shockley, 2003. Equilibrium "Anomalies", Journal of Finance 58:6, 2549-2580.

Foster, G., C. Olsen, and T. Shevlin, 1984. Earnings Releases, Anomalies and the Behavior of Security Returns, Accounting Review 59, 574-603.

Frankel, R., and C.M.C. Lee, 1998. Accounting valuation, market expectation, and cross-sectional stock returns. Journal of Accounting and Economics 25, 283-319.

Green, J., J. Hand, and F. Zhang, 2013. The Supraview of Return Predictive Signals, Review of Accounting Studies 18, 692-730.

Greig, A. C., 1992. Fundamental analysis and subsequent stock returns, Journal of Accounting and Economics 15:2, 413-442.

Grinblatt, M., and B. Han, 2005. Prospect Theory, Mental Accounting, and Momentum, Journal of Financial Economics 78, 311-339.

Grinblatt, M., and S. Titman, 1989. Mutual fund performance: An analysis of quarterly portfolio holdings, Journal of Business, 393-416.

Grinblatt, M., and S. Titman, 1992. The persistence of mutual fund performance, Journal of Finance 47, 1977-1984.

Grinblatt, M., and S. Titman, 1993. Performance measurement without benchmarks: An examination of mutual fund returns, Journal of Business 66, 47-68.

Grinblatt, M., and S. Titman, 1994. A study of monthly mutual fund returns and performance evaluation techniques, Journal of Financial and Quantitative Analysis 29, 419-444.

Harrison, M., and D. Kreps, 1979. Martingales and Arbitrage in Multiperiod Security Markets, Journal of Economic Theory 20, 381-408.

Harvey, C., Y. Liu, and H. Zhu, 2013. …and the Cross-Section of Expected Returns, Duke University Working Paper.

Hirshleifer, D., and D. Jiang, 2010. A Financing-Based Misvaluation Factor and the Cross-Section of Expected Returns, Review of Financial Studies 23:9, 3401-3436.

Hirshleifer, D., K. Hou, S.H. Teoh, and Y. Zhang, 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297-331.

Holthausen, R.W., and D.F. Larcker, 1992. The prediction of stock returns using financial statement information, Journal of Accounting and Economics 15:2, 373-411.

Hou, K., G.A. Karolyi, and B.-C. Kho, 2011. What Factors Drive Global Stock Returns?, Review of Financial Studies 24:8, 2527-2574.

Ikenberry, D., J. Lakonishok, and T. Vermaelen, 1995. Market underreaction to open market share repurchases, Journal of Financial Economics 39, 181-208.

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

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

Jegadeesh, N., and S. Titman, 2001. Profitability of momentum strategies: An evaluation of alternative explanations, Journal of Finance 56:2, 699-720.

Jiang, H., M. Verbeek, and Y. Wang, 2014. Information content when mutual funds deviate from benchmarks, Management Science, forthcoming.

Jones, C.P., and R.H. Litzenberger. 1970. Quarterly Earnings Reports and Intermediate Stock Price Trends, Journal of Finance 25, 143-148.

Joy, O.M., R.H. Litzenberger, and R.W. McEnally, 1977. The adjustment of stock prices to announcements of unanticipated changes in quarterly earnings, Journal of Accounting Research, 207-225.

Kogan, L., and M. Tian, 2013. Firm Characteristics and Empirical Factor Models: a Data-Mining Experiment. MIT Working Paper.

Lee, C.M.C., J. Myers, and B. Swaminathan, 1999. What is the Intrinsic Value of the Dow? Journal of Finance 54:5, 1693-1741.

Lev, B., and S.R. Thiagarajan, 1993. Fundamental information analysis, Journal of Accounting Research 31, 190-215.

Liu, J., D. Nissim, and J. Thomas. 2002. Equity valuation using multiples. Journal of Accounting Research 40:1, 135-172.

Livnat, J., and R. Mendenhall, 2006. Comparing the Post-Earnings Announcement Drift for Surprises Calculated from Analyst and Time-Series Forecasts, Journal of Accounting Research 44:1, 177-205.

Loughran, T., and J. Ritter, 1995. The new issues puzzle, Journal of Finance 50, 23–51.

Lyandres, E., L. Sun, and L. Zhang, 2008. The New Issues Puzzle: Testing the Investment-Based Explanation, Review of Financial Studies 21:6, 2827-2855.

Manaster, S., and R.J. Rendleman, 1982. Option Prices as Predictors of Equilibrium Stock Prices. Journal of Finance 37:4, 1043-1057.

McLean, D., and J. Pontiff, 2013. Does Academic Research Destroy Stock Return Predictability? Boston College working paper.

Merton, R.C., 1974. On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance 28, 449–470.

Michaely, R., H. Thaler, and L. Womack, 1995. Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift?, Journal of Finance 50:2, 573-608.

Mitchell, M.L., and E. Stafford, 2000. Managerial decisions and long-term stock price performance, Journal of Business 63, 411–433.

Mohanram, P.S., 2005. Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis, Review of Accounting Studies 10:2-3, 133-170.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium, Journal of Financial Economics 108, 1-28.

Ohlson, J., 1995. Earnings, book values, and dividends in equity valuation, Contemporary Accounting Research 11, 661-687.

Ou, J., and S. Penman, 1989. Financial statement analysis and the prediction of stock returns, Journal of Accounting and Economics 11, 295-330.

Pastor, L., and R.F. Stambaugh, 2003. Liquidity Risk and Expected Stock Returns, Journal of Political Economy 111:3, 642-685.

Penman, S.H., F. Reggiani, S.A. Richardson, and I. Tuna, 2014. An Accounting-Based Characteristic Model for Asset Pricing. London Business School Working Paper.

Perold, A.F., and B. Tierney, 1997. Numeric Investors L.P., HBS Case 9-298-012.

Piotroski, J.D., 2000. Value investing: The use of historical financial statement information to separate winners from losers, Journal of Accounting Research 38, 1-52.

Pontiff, J., and A. Woodgate, 2008. Share Issuance and Cross-Sectional Returns, Journal of Finance 93, 921-945.

Rendleman, R.J., C.P. Jones, and H.A. Latané, 1982. Empirical Anomalies based on Unexpected Earnings and the Importance of Risk Adjustments, Journal of Financial Economics 10, 269-287.

Rhodes-Kropf, M., D.T. Robinson, and S. Viswanathan, 2005. Valuation waves and merger activity: The empirical evidence. Journal of Financial Economics 77, 561-603.

Richardson, S.A., I. Tuna, and P. Wysocki, 2010. Accounting anomalies and fundamental analysis: A review of recent research advances, Journal of Accounting and Economics 50, 410-454.

Richardson, S.A., R.G. Sloan, M. Soliman, and I. Tuna, 2001. Information in Accruals about the Quality of Earnings, London Business School Working Paper.

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

Richardson, S.A., R.G. Sloan, M. Soliman, and I. Tuna, 2006. The Implications of Accounting Distortions and Growth for Accruals and Profitability, Accounting Review 81: 3, 713-743.

Ross, S.A., 1976. The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory 13, 341-360.

Ross, S.A., 1978. A Simple Approach to the Valuation of Risky Streams, Journal of Business 51:3, 453-475.

Schulmeister, S., 2009. Profitability of Technical Stock Trading: Has it Moved from Daily to Intraday Data?, Review of Financial Economics 18:4, 190-201.

Schultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs, Journal of Finance 58, 483–517.

Shumway, T., 1997. The Delisting Bias in CRSP Data, Journal of Finance 52:1, 327-340.

Sloan, R.G., 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? Accounting Review, 289-315.

Teoh, S.H., and T.J. Wong, 2002. Why do new issues and high accrual firms underperform? The role of analysts' credulity, Review of Financial Studies 15, 869-900.

Titman, S., K.C.J. Wei, and F. Xie, 2004. Capital Investments and Stock Returns, Journal of Financial and Quantitative Analysis 39:4, 677-700.

Titman, S., K.C.J. Wei, and F. Xie, 2013. Market Development and the Asset Growth Effect: International Evidence, Journal of Financial and Quantitative Analysis 48:5, 1405-1432.

Vassalou, M., and Y. Xing, 2004. Default Risk in Equity Returns, Journal of Finance 59:2, 831-868.

Wermers, R., 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction costs and Expenses, Journal of Finance 4, 1655-1695.

Figure 1: Signal Decay

The figure shows results from factor model time-series regressions. Stocks are sorted each month into quintiles based on the mispricing signal (*M*) and combined into equally-weighted portfolios. The signal is lagged between 0 and 35 months. A spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. The spread portfolio returns are regressed on a 6-factor model comprising the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev and LT_Rev. These factors are obtained from the Ken French data library. The 7-factor model additionally includes the PMU factor from the Robert Novy-Marx data library. All factors are adjusted to be in percent dimensions. The figure shows the alphas of time-series regressions of portfolio returns on the factors. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1980-12/2012. All variables are defined in Appendix B.

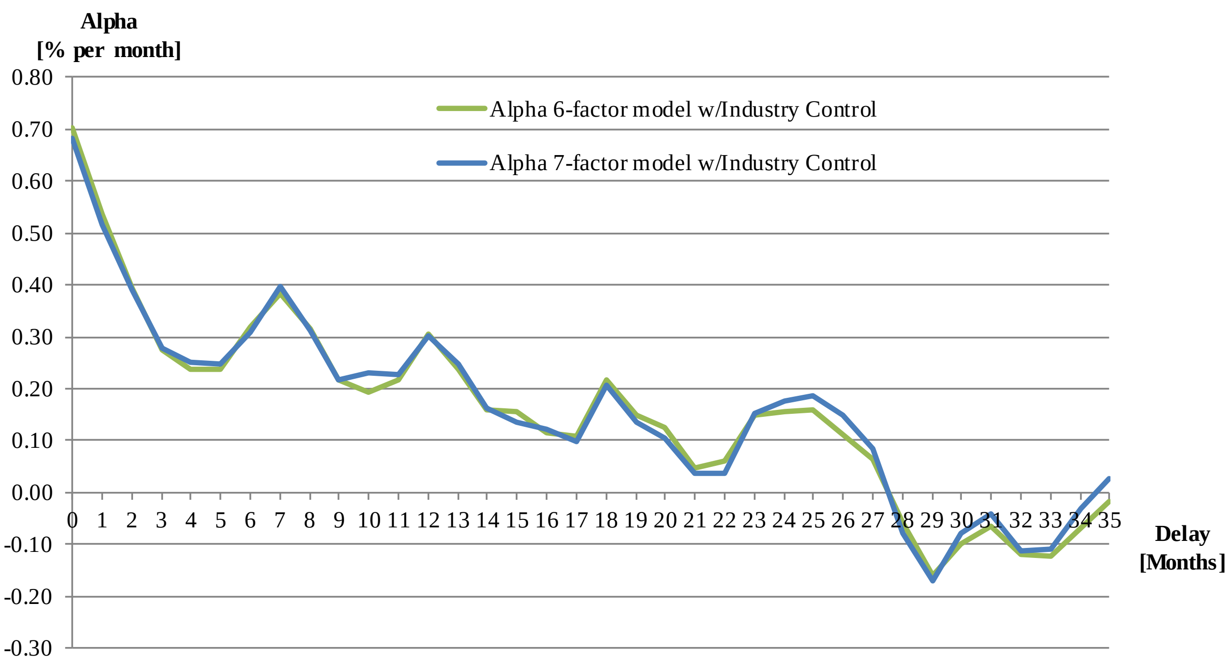

Figure 2: Accounting Weights

The figure shows results from factor model time-series regressions. Stocks are sorted each month into quintiles based on the mispricing signal (*M*) and combined into equally-weighted portfolios. The accounting weights are lagged between 0 and 11 months. A spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. The spread portfolio returns are regressed on a 6-factor model comprising the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev and LT_Rev. These factors are obtained from the Ken French data library. The 7-factor model additionally includes the PMU factor from the Robert Novy-Marx data library. All factors are adjusted to be in percent dimensions. The figure shows the alphas of time-series regressions of portfolio returns on the factors. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1978-12/2012. All variables are defined in Appendix B.

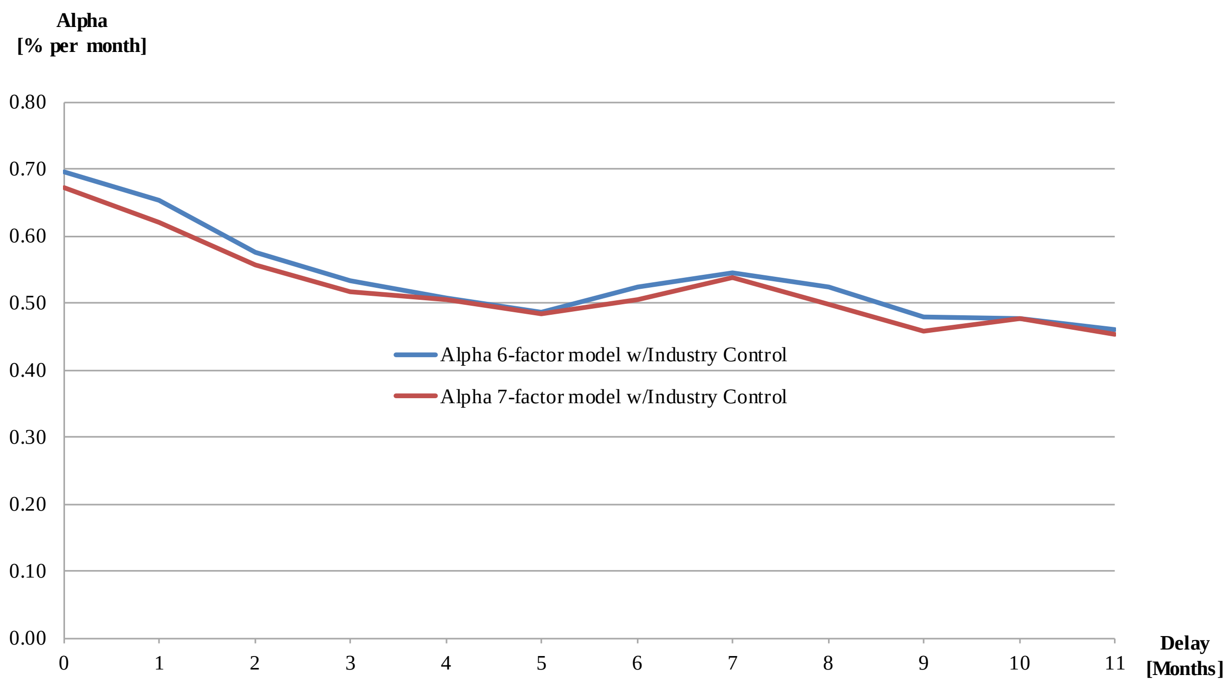

Table 1: Summary Statistics by Mispricing Signal Quintiles

The table reports averages of a number of characteristics of portfolios firms. In particular, the table reports the time-series average of the mean characteristics across all firms ("All"), the average cross-sectional correlation of the characteristic with the mispricing signal *M* ("Correlation"), as well as the average of the mean characteristics across quintiles of firms sorted by the mispricing signal *M* from Q1 (most overpriced) to Q5 (most underpriced). The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

			Signal Quintiles				
	All	Correlation	Q1 (Overvalued)	Q2	Q3	Q4	Q5 (Undervalued)
Mispricing Signal (M)	0.5406	1.000	-2.2251	-0.3213	0.2820	1.0265	3.9403
Market Capitalization	2,326.9	-0.068	2,020.6	4,338.6	3,283.1	1531.7	458.6
Book/Market	0.6736	0.262	0.5666	0.4877	0.5802	0.7203	1.0135
Beta	0.9472	-0.116	1.0112	1.0132	0.9908	0.9306	0.7880
Accruals	0.1445	-0.029	0.1407	0.1678	0.1585	0.1424	0.1164
SUE	0.0010	0.029	-0.0006	0.0014	0.0015	0.0013	0.0012
Gross Profitability	0.3911	0.035	0.3592	0.4040	0.4046	0.3957	0.3918
Default Risk	0.0151	0.033	0.0232	0.0073	0.0064	0.0096	0.0292
Prior Month Return t	2.0892	-0.028	3.1322	2.7201	2.1204	1.4727	1.0015
Return from Month t-1 to t-11	23.439	-0.040	31.077	30.420	24.352	17.502	13.892
Return from Month t-12 to t-59	108.88	-0.040	114.84	125.56	120.37	102.97	80.068

Table 2: Stock Returns and Mispricing Signal Quintiles

The table reports averages and selected test statistics of portfolio returns. In particular, the table reports the time-series average of the mean return across all firms ("All"), the average cross-sectional correlation between returns and the mispricing signal *M* ("Correlation"), as well as the average return across quintiles of firms sorted by the mispricing signal *M* from Q1 (most overpriced) to Q5 (most underpriced). The table also shows the time-series average of the quintile spread (the difference between the return for the most undervalued firms (5th quintile) and the most overvalued firms (1st quintile)) as well as the associated *t*-statistic of a test against 0. Moreover, the table reports the fraction of time-series observations of the quintile spread that is greater than zero and the *p*-value of a binomial test of a test against 50%. Returns are either equally-weighted or value-weighted Panel A reports results for equal weighted portfolios of the characteristic, while Panel B shows results for value weighted portfolios of the characteristic. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

			Signal Quintiles					Q5-Q1 (Undervalued - Overvalued)			
	All	Correlation	Q1 (Overvalued)	Q2	Q3	Q4	Q5 (Undervalued)	Fraction > 0	*p-value*	Average	*t-stat*
Panel A: Equally-weighted Portfolios											
Return in Month *t*+1	1.2603	0.0159	0.9259	1.0989	1.2748	1.3983	1.6039	61.1	[0.00]	0.6781	[4.85]
Return in Month *t*+1 (1977-1988)	1.6280	0.0195	1.3151	1.4005	1.5740	1.8172	2.0339	61.8	[0.00]	0.7188	[3.72]
Return in Month *t*+1 (1989-2000)	1.2095	0.0150	0.7613	1.2274	1.4251	1.3144	1.3190	56.9	[0.10]	0.5577	[1.76]
Return in Month *t*+1 (2001-2012)	0.9434	0.0132	0.7012	0.6687	0.8255	1.0633	1.4589	64.6	[0.00]	0.7577	[3.82]
Panel B: Value-weighted Portfolios											
Return in Month *t*+1	1.0105	0.0142	0.9994	0.8564	1.1342	1.3092	1.4416	56.0	[0.01]	0.4422	[2.63]
Return in Month *t*+1 (1977-1988)	1.2095	0.0194	1.2138	1.0947	1.3504	1.6746	1.9487	59.7	[0.02]	0.7349	[2.52]
Return in Month *t*+1 (1989-2000)	1.3437	-0.0011	1.1440	1.1446	1.4007	1.2875	1.3585	52.8	[0.50]	0.2145	[0.71]
Return in Month *t*+1 (2001-2012)	0.4784	0.0241	0.6403	0.3299	0.6515	0.9656	1.0176	55.6	[0.18]	0.3773	[1.33]

Table 3: Fama-MacBeth Cross-Sectional Regressions

The table shows results from Fama MacBeth (1973) regressions. Across different specifications, the return in the next period is regressed on the mispricing signal *M*, market beta, book-to-market, market capitalization, short-term reversal, momentum, long-term reversal, accruals, SUE, gross profitability, and default risk. Panel A shows results for regressions with regular variables (in some cases with standard logarithmic transformations). Panel B shows results using quintile dummies for independent variables. Quintiles are determined in each month based on independent sorts of the sample firms that have non-missing values for all characteristics. Size quintiles are based on NYSE breakpoints for market. The regressions include dummy variables for quintiles 2, 3, 4 and 5 of each characteristic, but for brevity the table only displays the coefficients of selected regressors. Panel C shows results for the same specifications as Panel B but limited to the period 1993-2012. All regressions use dummy variables based on the 38 Fama French industry classifications. The table shows the average regression coefficients, associated *t*-statistics, as well as the average number of observations and adjusted R-Squared. *, **, and *** indicate statistical significance at the 10% (5%, 1%) significance level. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

Panel A: Regressions with Regular Variables

	(1)			(2)			(3)			(4)			(5)			(6)			(7)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Mispricing Signal (*M*)	0.0903	[5.95]	***				0.0616	[4.42]	***				0.0826	[5.44]	***				0.0448	[3.24]	***
Beta				-0.0545	[-0.41]		-0.0475	[-0.36]								0.0142	[0.11]		0.0131	[0.10]	
Log(Book/Market)				0.2710	[4.19]	***	0.2267	[3.44]	***							0.3580	[5.23]	***	0.3197	[4.58]	***
Log(Market Capitalization)				-0.0384	[-1.01]		-0.0417	[-1.05]								-0.0355	[-0.94]		-0.0459	[-1.16]	
Short-term Reversal				-0.0364	-10.16]	***	-0.0363	-10.16]	***							-0.0412	-11.53]	***	-0.0410	-11.53]	***
Momentum				0.0055	[4.87]	***	0.0055	[4.84]	***							0.0032	[3.15]	***	0.0033	[3.21]	***
Long-term Reversal				-0.0004	[-2.39]	**	-0.0004	[-2.46]	**							0.0000	[-0.03]		0.0000	[-0.04]	
Accruals										-1.1569	[-8.61]	***	-1.1482	[-8.72]	***	-1.0514	-10.75]	***	-1.0686	-10.95]	***
SUE										15.963	[12.42]	***	15.812	[12.23]	***	16.063	[14.20]	***	15.786	[14.01]	***
Gross Profitability										0.9812	[6.90]	***	0.9812	[6.90]	***	1.1916	[8.13]	***	1.1634	[7.91]	***
Default Risk										-1.3034	[-2.00]	**	-1.3368	[-2.04]	**	-2.3186	[-4.16]	***	-2.1970	[-3.94]	***
Intercept	1.0568	[3.73]	***	1.3600	[4.08]	***	1.3360	[3.81]	***	0.7541	[2.67]	***	0.7210	[2.53]	**	0.9683	[2.80]	***	1.0163	[2.81]	***
Observations	1,219			1,219			1,219			1,219			1,219			1,219			1,219		
Adj. RSquare	0.042			0.076			0.078			0.049			0.053			0.083			0.084		
Industry Control	Yes			Yes			Yes			Yes			Yes			Yes			Yes		

(continued)

Table 3: Fama-MacBeth Cross-Sectional Regressions (continued)

Panel B: Regressions with Quintile Dummies

	(1)			(2)			(3)			(4)			(5)			(6)			(7)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Mispricing Signal (*M*) Q5	0.7292	[5.66]	***				0.5326	[5.75]	***				0.7216	[5.64]	***				0.3849	[3.98]	***
Beta Q5				-0.1360	[-0.79]		-0.1428	[-0.84]								-0.0634	[-0.40]		-0.0874	[-0.55]	
Market Capitalization Q5				-0.2177	[-1.25]		-0.2125	[-1.27]								-0.2072	[-1.26]		-0.2260	[-1.41]	
Book/Market Q5				0.4742	[3.54]	***	0.2656	[2.03]	**							0.6894	[4.89]	***	0.5193	[3.69]	***
Short-term Reversal Q5				-1.3278	[-9.62]	***	-1.3151	[-9.58]	***							-1.5540	-11.64]	***	-1.5400	-11.60]	***
Momentum Q5				0.9517	[5.70]	***	0.9292	[5.63]	***							0.3713	[2.39]	**	0.3695	[2.40]	**
Long-term Reversal Q5				-0.1864	[-2.05]	**	-0.2364	[-2.63]	***							0.0769	[0.91]		0.0543	[0.65]	
Accruals Q5										-0.7631	[-8.39]	***	-0.7569	[-8.56]	***	-0.7270	-10.35]	***	-0.7494	-10.66]	***
SUE Q5										1.1591	[11.99]	***	1.1689	[12.09]	***	1.2372	[15.18]	***	1.2225	[14.94]	***
Gross Profitability Q5										0.6275	[6.57]	***	0.6016	[6.42]	***	0.6678	[7.08]	***	0.6234	[6.61]	***
Default Risk Q5										0.0605	[0.44]		-0.0365	[-0.27]		-0.2611	[-2.51]	**	-0.2452	[-2.35]	**
Intercept	0.7690	[2.60]	***	1.1974	[3.23]	***	1.1211	[2.97]	***	0.4334	[1.44]		0.1954	[0.63]		0.7616	[2.06]	**	0.8011	[2.13]	**
Observations	1,219			1,219			1,219			1,219			1,219			1,219			1,219		
Adj. RSquare	0.045			0.078			0.079			0.056			0.061			0.085			0.086		
Industry Control	Yes			Yes			Yes			Yes			Yes			Yes			Yes		

(continued)

Table 3: Fama-MacBeth Cross-Sectional Regressions (continued)

Panel C: Regressions with Quintile Dummies for 1993-2012

	(1)			(2)			(3)			(4)			(5)			(6)			(7)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Mispricing Signal (*M*) Q5	0.7866	[4.23]	***				0.6519	[4.88]	***				0.8065	[4.33]	***				0.5412	[3.91]	***
Beta Q5				-0.1672	[-0.64]		-0.1719	[-0.67]								-0.1210	[-0.50]		-0.1492	[-0.63]	
Market Capitalization Q5				-0.0608	[-0.25]		-0.1770	[-0.75]								-0.0104	[-0.04]		-0.1204	[-0.52]	
Book/Market Q5				0.4818	[2.35]	**	0.2205	[1.12]								0.6999	[3.26]	***	0.4655	[2.20]	**
Short-term Reversal Q5				-1.1009	[-5.08]	***	-1.0877	[-5.06]	***							-1.2632	[-6.09]	***	-1.2449	[-6.05]	***
Momentum Q5				0.6837	[2.60]	***	0.6638	[2.57]	**							0.3237	[1.34]		0.3357	[1.40]	
Long-term Reversal Q5				-0.3162	[-2.42]	**	-0.3725	[-2.92]	***							-0.0635	[-0.53]		-0.0839	[-0.71]	
Accruals Q5										-0.8061	[-6.15]	***	-0.8185	[-6.46]	***	-0.7229	[-7.35]	***	-0.7611	[-7.71]	***
SUE Q5										0.8135	[5.92]	***	0.8324	[6.09]	***	0.8741	[7.67]	***	0.8474	[7.42]	***
Gross Profitability Q5										0.7058	[4.91]	***	0.6900	[4.97]	***	0.7963	[5.65]	***	0.7301	[5.20]	***
Default Risk Q5										0.0715	[0.36]		0.0157	[0.08]		-0.1260	[-0.85]		-0.1035	[-0.70]	
Intercept	0.3208	[0.78]		0.8730	[1.58]		0.8367	[1.48]		0.2579	[0.56]		-0.0099	[-0.02]		0.4171	[0.74]		0.4939	[0.86]	
Observations	1,397			1,397			1,397			1,397			1,397			1,397			1,397		
Adj. RSquare	0.047			0.080			0.081			0.058			0.063			0.087			0.088		
Industry Control	Yes			Yes			Yes			Yes			Yes			Yes			Yes		

Table 4: Factor Model Time-Series Regressions

The table shows results from factor model time-series regressions. Stocks are sorted each month into quintiles based on the mispricing signal (*M*) and combined into equally-weighted (Panel A) or value-weighted (Panel B) portfolios. Regressions are performed separately for each of the portfolios, where the portfolio of the most overvalued stocks is Q1, while the most undervalued stocks are in portfolio Q5. Additionally, a spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. Portfolio returns are regressed on the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, LT_Rev, and PMU. These factors are obtained from the data libraries from Ken French and Robert Novy-Marx, respectively, and adjusted to be in percent dimensions. The table reports the regression coefficients and associated *t*-statistics of time-series regressions of portfolio returns (in excess of the industry portfolios based on the 38 Fama French industry classifications) on the factors. The table shows the regression coefficients, associated *t*-statistics, as well as the number of observations and R-Squared. *, **, and *** indicate statistical significance at the 10% (5%, 1%) significance level. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

(continued)

Table 4: Factor Model Time-Series Regressions (continued)

Panel A: Equal-weighted Portfolios

	Q1 (Overvalued)		Q2		Q3		Q4		Q5 (Undervalued)		Q5-Q1 (Undervalued - Overvalued)	
	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat
Industry-Adjusted Return	-0.3485	[-3.61] ***	-0.2245	[-1.91] *	-0.0373	[-0.35]	0.1091	[1.01]	0.3331	[3.51] ***	0.6815	[5.85] ***
Alpha	-0.4500	[-4.94] ***	-0.2615	[-2.47] **	-0.0975	[-0.99]	0.0312	[0.31]	0.2639	[3.09] ***	0.7139	[6.83] ***
Mkt_RF	0.0424	[1.97] **	0.0313	[1.25]	0.0375	[1.62]	0.0277	[1.17]	0.0162	[0.80]	-0.0262	[-1.06]
SMB	-0.1757	[-5.25] ***	-0.2708	[-6.97] ***	-0.2348	[-6.52] ***	-0.2158	[-5.88] ***	-0.0827	[-2.64] ***	0.0931	[2.42] **
HML	-0.0097	[-0.27]	-0.1560	[-3.73] ***	-0.0028	[-0.07]	0.2119	[5.36] ***	0.3233	[9.57] ***	0.3330	[8.05] ***
Mom	0.1753	[8.63] ***	0.2066	[8.76] ***	0.1803	[8.25] ***	0.1092	[4.90] ***	0.0120	[0.63]	-0.1634	[-7.01] ***
ST_Rev	-0.0110	[-0.40]	-0.0569	[-1.77] *	-0.0606	[-2.03] **	0.0084	[0.28]	0.0180	[0.69]	0.0290	[0.91]
LT_Rev	0.0400	[0.98]	0.0711	[1.50]	-0.0126	[-0.29]	-0.1305	[-2.90] ***	-0.1643	[-4.28] ***	-0.2042	[-4.35] ***
RSquare	0.20		0.27		0.24		0.24		0.27		0.28	
Observations	432		432		432		432		432		432	
Alpha	-0.5355	[-5.82] ***	-0.3423	[-3.18] ***	-0.2158	[-2.21] **	-0.1236	[-1.27]	0.1561	[1.84] *	0.6916	[6.44] ***
Mkt_RF	0.0502	[2.37] **	0.0387	[1.56]	0.0484	[2.15] **	0.0419	[1.86] *	0.0260	[1.33]	-0.0242	[-0.98]
SMB	-0.1580	[-4.77] ***	-0.2541	[-6.56] ***	-0.2103	[-5.97] ***	-0.1837	[-5.23] ***	-0.0603	[-1.97] **	0.0977	[2.52] **
HML	0.0384	[1.03]	-0.1107	[-2.53] **	0.0636	[1.60]	0.2988	[7.55] ***	0.3838	[11.14] ***	0.3455	[7.92] ***
Mom	0.1759	[8.82] ***	0.2071	[8.88] ***	0.1811	[8.54] ***	0.1103	[5.22] ***	0.0127	[0.69]	-0.1632	[-7.00] ***
ST_Rev	-0.0218	[-0.80]	-0.0671	[-2.10] **	-0.0756	[-2.60] ***	-0.0112	[-0.39]	0.0044	[0.17]	0.0262	[0.82]
LT_Rev	0.0171	[0.42]	0.0495	[1.04]	-0.0442	[-1.02]	-0.1719	[-3.99] ***	-0.1931	[-5.15] ***	-0.2102	[-4.43] ***
PMU	0.1589	[4.05] ***	0.1500	[3.27] ***	0.2197	[5.27] ***	0.2875	[6.91] ***	0.2002	[5.53] ***	0.0414	[0.90]
RSquare	0.23		0.29		0.29		0.31		0.32		0.28	
Observations	432		432		432		432		432		432	

(continued)

Table 4: Factor Model Time-Series Regressions (continued)

Panel B: Value-weighted Portfolios

	Q1 (Overvalued)			Q2			Q3			Q4			Q5 (Undervalued)			Q5-Q1 (Undervalued - Overvalued)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Industry-Adjusted Return	-0.2773	[-1.85]	*	-0.42271	[-2.44]	**	-0.2191	[-1.30]		0.02479	[0.17]		0.22593	[1.86]	*	0.50318	[3.35]	***
Alpha	-0.1737	[-1.34]		-0.1902	[-1.58]		-0.0845	[-0.74]		0.0442	[0.39]		0.1480	[1.33]		0.3216	[2.19]	**
Mkt_RF	0.0957	[3.13]	***	-0.0272	[-0.96]		-0.0333	[-1.23]		0.0410	[1.52]		0.1139	[4.35]	***	0.0182	[0.52]	
SMB	-0.5796	[-12.19]	***	-0.8892	[-20.13]	***	-0.8457	[-20.08]	***	-0.6228	[-14.88]	***	-0.3162	[-7.76]	***	0.2635	[4.89]	***
HML	-0.0914	[-1.78]	*	-0.2384	[-5.01]	***	-0.1215	[-2.68]	***	0.0972	[2.15]	**	0.2310	[5.26]	***	0.3224	[5.55]	***
Mom	0.0890	[3.08]	***	0.1523	[5.68]	***	0.2050	[8.02]	***	0.1064	[4.19]	***	0.0424	[1.72]	*	-0.0466	[-1.42]	
ST_Rev	-0.1268	[-3.22]	***	-0.1483	[-4.06]	***	-0.0607	[-1.74]	*	0.0637	[1.84]	*	0.0327	[0.97]		0.1596	[3.57]	***
LT_Rev	0.0232	[0.40]		0.1553	[2.87]	***	0.0833	[1.62]		-0.0494	[-0.96]		-0.0828	[-1.66]	*	-0.1060	[-1.61]	
RSquare	0.33			0.57			0.58			0.44			0.25			0.15		
Observations	432			432			432			432			432			432		
Alpha	-0.2983	[-2.29]	**	-0.2864	[-2.35]	**	-0.2724	[-2.45]	**	-0.1451	[-1.32]		0.0249	[0.22]		0.3233	[2.14]	**
Mkt_RF	0.1072	[3.56]	***	-0.0183	[-0.65]		-0.0160	[-0.63]		0.0583	[2.30]	**	0.1252	[4.89]	***	0.0180	[0.52]	
SMB	-0.5538	[-11.77]	***	-0.8693	[-19.75]	***	-0.8068	[-20.14]	***	-0.5836	[-14.69]	***	-0.2907	[-7.26]	***	0.2631	[4.83]	***
HML	-0.0214	[-0.40]		-0.1845	[-3.72]	***	-0.0160	[-0.35]		0.2035	[4.55]	***	0.3001	[6.65]	***	0.3215	[5.24]	***
Mom	0.0899	[3.17]	***	0.1529	[5.77]	***	0.2063	[8.56]	***	0.1077	[4.50]	***	0.0432	[1.79]	*	-0.0466	[-1.42]	
ST_Rev	-0.1426	[-3.67]	***	-0.1605	[-4.42]	***	-0.0845	[-2.56]	**	0.0397	[1.21]		0.0172	[0.52]		0.1598	[3.56]	***
LT_Rev	-0.0101	[-0.18]		0.1296	[2.40]	**	0.0330	[0.67]		-0.1001	[-2.06]	**	-0.1157	[-2.36]	**	-0.1056	[-1.58]	
PMU	0.2315	[4.16]	***	0.1786	[3.43]	***	0.3490	[7.36]	***	0.3516	[7.47]	***	0.2285	[4.82]	***	-0.0031	[-0.05]	
RSquare	0.35			0.58			0.63			0.51			0.29			0.15		
Observations	432			432			432			432			432			432		

Table 5: Double-Sorted Portfolios

The table shows results from sorting stocks by size, value and mispricing signal. At the end of each month, stocks are sorted sequentially into quintiles by NYSE market capitalization (Panel A) or book-to-market (Panel B) and then by mispricing signal (*M*). The table reports various statistics for the equally-weighted portfolios of the most over- (Q1) and under-valued firms (Q5) and the spread between the two (Q5-Q1), including the time-series average of the industry-adjusted return, and alphas from time-series regressions of portfolio returns (in excess of the industry portfolios based on the 38 Fama French industry classifications) on a 6-factor model (with the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, LT_Rev) and 7-factor model (that adds PMU to the 6-factor model). These factors are obtained from the data libraries from Ken French and Robert Novy-Marx, respectively, and adjusted to be in percent dimensions. The table also reports of time series average of equal firm weightings of percentile and quintile ranks. Each month's distribution of quintile ranks generate weights that are applied to corresponding performance measures in Table 4 Panel A. The time series average of weighted Table 4 Panel A performance is reported in the "Predicted" column. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

Panel A: Size Sorts

Market Capitalization	Signal	Percent Rank Mean	Quintile Rank Mean	Industry-Adjusted Returns: Mean	*t*-stat	Predicted	Difference	*t*-stat	6-Factor Model: Alpha	*t*-stat	Predicted	Difference	*t*-stat	7-Factor Model: Alpha	*t*-stat	Predicted	Difference	*t*-stat
	Q1	0.19	1.59	-0.3512	[-3.41]	-0.2854	-0.0658	[-0.85]	-0.5372	[-5.37]	-0.3712	-0.1660	[-2.36]	-0.6302	[-6.24]	-0.4631	-0.1672	[-2.38]
Small	Q5	0.93	4.95	0.4193	[4.29]	0.3044	0.1149	[2.06]	0.3405	[3.94]	0.2311	0.1094	[2.11]	0.2679	[3.06]	0.1221	0.1458	[2.85]
	Q5-Q1	0.74	3.36	0.7705	[5.95]	0.5898	0.1808	[1.88]	0.8777	[7.03]	0.6024	0.2754	[3.04]	0.8982	[7.00]	0.5852	0.3130	[3.48]
	Q1	0.12	1.19	-0.4009	[-2.80]	-0.2811	-0.1198	[-1.21]	-0.6438	[-4.86]	-0.3703	-0.2735	[-3.19]	-0.6505	[-4.77]	-0.4550	-0.1954	[-2.29]
Q2	Q5	0.79	4.40	0.3474	[2.68]	0.2900	0.0574	[0.74]	0.2554	[2.00]	0.2015	0.0539	[0.71]	0.1358	[1.05]	0.0703	0.0655	[0.87]
	Q5-Q1	0.66	3.22	0.7482	[4.60]	0.5711	0.1771	[1.51]	0.8992	[6.51]	0.5718	0.3274	[2.81]	0.7863	[5.62]	0.5254	0.2609	[2.25]
	Q1	0.13	1.22	-0.2469	[-1.64]	-0.2931	0.0462	[0.49]	-0.3829	[-2.71]	-0.3728	-0.0102	[-0.11]	-0.3994	[-2.74]	-0.4562	0.0567	[0.63]
Q3	Q5	0.74	4.18	0.2729	[1.95]	0.2631	0.0099	[0.12]	0.2447	[1.92]	0.1776	0.0671	[0.91]	0.0783	[0.62]	0.0476	0.0306	[0.42]
	Q5-Q1	0.61	2.96	0.5198	[2.93]	0.5561	-0.0363	[-0.30]	0.6276	[4.31]	0.5503	0.0773	[0.68]	0.4777	[3.26]	0.5038	-0.0261	[-0.23]
	Q1	0.13	1.26	-0.2556	[-1.58]	-0.2903	0.0347	[0.33]	-0.2038	[-1.43]	-0.3670	0.1632	[1.70]	-0.2373	[-1.62]	-0.4519	0.2147	[2.24]
Q4	Q5	0.70	3.99	0.1569	[1.05]	0.2275	-0.0705	[-0.72]	0.2123	[1.66]	0.1507	0.0616	[0.74]	0.0628	[0.49]	0.0228	0.0400	[0.48]
	Q5-Q1	0.57	2.73	0.4125	[2.70]	0.5178	-0.1053	[-0.86]	0.4161	[3.03]	0.5176	-0.1015	[-0.85]	0.3001	[2.16]	0.4748	-0.1747	[-1.46]
	Q1	0.16	1.32	-0.4256	[-2.42]	-0.2919	-0.1337	[-1.11]	-0.2761	[-2.13]	-0.3616	0.0856	[0.96]	-0.3398	[-2.57]	-0.4478	0.1080	[1.21]
Large	Q5	0.62	3.62	-0.0178	[-0.11]	0.1306	-0.1484	[-1.42]	0.1381	[1.29]	0.0643	0.0738	[1.02]	0.0270	[0.25]	-0.0617	0.0888	[1.22]
	Q5-Q1	0.47	2.29	0.4077	[2.76]	0.4225	-0.0147	[-0.13]	0.4141	[3.24]	0.4259	-0.0118	[-0.10]	0.3668	[2.79]	0.3861	-0.0192	[-0.17]

(continued)

Table 5: Double-Sorted Portfolios (continued)

Panel B: Value Sorts

Book/ Market	Signal	Percent Rank Mean	Quintile Rank Mean	Industry-Adjusted Returns Mean	t-stat	Predicted	Difference	t-stat	6-Factor Model Alpha	t-stat	Predicted	Difference	t-stat	7-Factor Model Alpha	t-stat	Predicted	Difference	t-stat
	Q1	0.08	1.02	-0.4129	[-2.54]	-0.3316	-0.0813	[-0.66]	-0.4559	[-3.22]	-0.4349	-0.0211	[-0.21]	-0.4395	[-3.02]	-0.5204	0.0809	[0.82]
Growth	Q5	0.65	3.73	-0.0457	[-0.31]	0.0061	-0.0518	[-0.51]	-0.2023	[-1.51]	-0.0219	-0.1804	[-1.93]	-0.3961	[-3.00]	-0.1454	-0.2507	[-2.71]
	Q5-Q1	0.57	2.71	0.3672	[1.82]	0.3377	0.0295	[0.19]	0.2536	[1.32]	0.4130	-0.1593	[-1.07]	0.0434	[0.23]	0.3750	-0.3316	[-2.29]
	Q1	0.13	1.15	-0.3045	[-2.66]	-0.3167	0.0121	[0.16]	-0.3629	[-3.25]	-0.4097	0.0468	[0.65]	-0.4017	[-3.50]	-0.4946	0.0930	[1.30]
Q2	Q5	0.78	4.36	0.0008	[0.01]	0.1782	-0.1774	[-2.16]	-0.0725	[-0.57]	0.1242	-0.1967	[-2.61]	-0.2307	[-1.83]	-0.0137	-0.2171	[-2.88]
	Q5-Q1	0.65	3.22	0.3054	[1.83]	0.4949	-0.1895	[-1.67]	0.2904	[1.77]	0.5338	-0.2434	[-2.24]	0.1709	[1.02]	0.4810	-0.3100	[-2.88]
	Q1	0.15	1.33	-0.2843	[-2.50]	-0.3181	0.0338	[0.48]	-0.4476	[-4.21]	-0.4008	-0.0469	[-0.68]	-0.5461	[-5.08]	-0.4845	-0.0616	[-0.89]
Q3	Q5	0.85	4.72	0.2163	[1.88]	0.2752	-0.0589	[-0.86]	0.1450	[1.30]	0.2043	-0.0594	[-0.91]	0.0089	[0.08]	0.0818	-0.0729	[-1.12]
	Q5-Q1	0.70	3.39	0.5007	[3.63]	0.5933	-0.0927	[-0.94]	0.5926	[4.32]	0.6051	-0.0125	[-0.13]	0.5550	[3.94]	0.5663	-0.0113	[-0.12]
	Q1	0.16	1.37	-0.2379	[-2.00]	-0.3017	0.0638	[0.78]	-0.4020	[-3.58]	-0.3859	-0.0161	[-0.23]	-0.5143	[-4.56]	-0.4711	-0.0432	[-0.61]
Q4	Q5	0.91	5.00	0.4158	[3.63]	0.3329	0.0829	[1.24]	0.3793	[3.44]	0.2638	0.1155	[1.74]	0.2969	[2.65]	0.1560	0.1410	[2.13]
	Q5-Q1	0.76	3.63	0.6538	[4.79]	0.6346	0.0191	[0.17]	0.7813	[5.64]	0.6496	0.1316	[1.32]	0.8112	[5.70]	0.6271	0.1842	[1.86]
	Q1	0.16	1.45	-0.0860	[-0.70]	-0.2743	0.1883	[1.71]	-0.2467	[-2.14]	-0.3676	0.1209	[1.38]	-0.3299	[-2.82]	-0.4571	0.1272	[1.45]
Value	Q5	0.97	5.00	0.5024	[4.19]	0.3331	0.1693	[2.20]	0.5333	[4.88]	0.2639	0.2695	[3.78]	0.4736	[4.24]	0.1561	0.3175	[4.49]
	Q5-Q1	0.81	3.55	0.5884	[3.97]	0.6074	-0.0189	[-0.15]	0.7801	[5.18]	0.6315	0.1486	[1.31]	0.8035	[5.19]	0.6132	0.1903	[1.67]

Table 6: Risk Analysis

The table shows results from analysis of the risk of the mispricing signal strategy. In particular, it shows results for a pseudo mispricing signal that only uses a constant in the mispricing signal regression. Stocks are sorted each month into quintiles based on this pseudo mispricing signal and combined into equally-weighted portfolios. Factor model time-series regressions are performed separately for each of the portfolios, where the portfolio of the most overvalued stocks is Q1, while the most undervalued stocks are in portfolio Q5. Additionally, a spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. Portfolio returns are regressed on the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, LT_Rev and PMU. These factors are obtained from the data libraries from Ken French and Robert Novy-Marx, respectively, and adjusted to be in percent dimensions. The table reports the regression coefficients and associated *t*-statistics of time-series regressions of portfolio returns (in excess of the industry portfolios based on the 38 Fama French industry classifications) on the factors. The table shows the regression coefficients, associated *t*-statistics, as well as the number of observations and R-Squared. *, **, and *** indicate statistical significance at the 10% (5%, 1%) significance level. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1977-12/2012. All variables are defined in Appendix B.

(continued)

Table 6: Risk Analysis (continued)

	Q1 (Overvalued)		Q2		Q3		Q4		Q5 (Undervalued)		Q5-Q1 (Undervalued - Overvalued)	
	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat	Coef	*t*-stat
Industry-Adjusted Return	-0.2153	[-1.47]	-0.1085	[-0.93]	0.0245	[0.24]	0.0938	[1.12]	0.0373	[0.41]	0.2526	[1.74] *
Intercept	-0.1052	[-1.04]	-0.1509	[-1.36]	-0.1112	[-1.07]	-0.0730	[-0.88]	-0.0738	[-0.99]	0.0314	[0.30]
Mkt_RF	0.0560	[2.36] **	0.0926	[3.53] ***	0.0886	[3.63] ***	0.0335	[1.72] *	-0.1158	[-6.62] ***	-0.1718	[-6.91] ***
SMB	-0.7585	-20.50] ***	-0.2903	[-7.12] ***	0.0049	[0.13]	0.0664	[2.18] **	-0.0027	[-0.10]	0.7558	[19.53] ***
HML	-0.0902	[-2.26] **	0.0108	[0.25]	0.0386	[0.94]	0.1545	[4.72] ***	0.2532	[8.63] ***	0.3434	[8.24] ***
Mom	0.1429	[6.36] ***	0.1238	[5.00] ***	0.1241	[5.38] ***	0.1257	[6.81] ***	0.1669	[10.10] ***	0.0240	[1.02]
ST_Rev	-0.0616	[-2.01] **	-0.0536	[-1.59]	-0.0129	[-0.41]	0.0186	[0.74]	0.0073	[0.32]	0.0689	[2.15] **
LT_Rev	0.0199	[0.44]	-0.0198	[-0.40]	-0.0380	[-0.82]	-0.0520	[-1.40]	-0.1063	[-3.19] ***	-0.1262	[-2.66] ***
RSquare	0.57		0.19		0.09		0.13		0.41		0.53	
Observations	432		432		432		432		432		432	
Intercept	-0.2339	[-2.34] **	-0.2842	[-2.57] **	-0.2270	[-2.19] **	-0.1738	[-2.11] **	-0.1423	[-1.90] *	0.0916	[0.85]
Mkt_RF	0.0678	[2.94] ***	0.1048	[4.11] ***	0.0993	[4.16] ***	0.0428	[2.25] **	-0.1095	[-6.34] ***	-0.1774	[-7.14] ***
SMB	-0.7318	-20.29] ***	-0.2627	[-6.58] ***	0.0288	[0.77]	0.0873	[2.94] ***	0.0115	[0.43]	0.7433	[19.15] ***
HML	-0.0179	[-0.44]	0.0856	[1.90] *	0.1036	[2.46] **	0.2111	[6.30] ***	0.2917	[9.59] ***	0.3096	[7.08] ***
Mom	0.1438	[6.62] ***	0.1247	[5.19] ***	0.1249	[5.56] ***	0.1264	[7.07] ***	0.1674	[10.30] ***	0.0236	[1.01]
ST_Rev	-0.0779	[-2.62] ***	-0.0705	[-2.14] **	-0.0276	[-0.90]	0.0058	[0.24]	-0.0014	[-0.06]	0.0765	[2.39] **
LT_Rev	-0.0146	[-0.33]	-0.0555	[-1.13]	-0.0690	[-1.51]	-0.0789	[-2.17] **	-0.1247	[-3.77] ***	-0.1101	[-2.31] **
PMU	0.2390	[5.60] ***	0.2475	[5.24] ***	0.2151	[4.86] ***	0.1871	[5.31] ***	0.1273	[3.98] ***	-0.1117	[-2.43] **
RSquare	0.60		0.24		0.14		0.18		0.43		0.53	
Observations	432		432		432		432		432		432	

Table 7: Buy-and-Hold Returns

The table shows results from factor model time-series regressions for buy-and-hold returns. Stocks are sorted each month into quintiles based on the mispricing signal (*M*) and combined into equally-weighted portfolios. Following Jegadeesh and Titman (1993, 2001) each portfolio is held for 12 months. The strategy return is the simple average of the returns to the twelve overlapping portfolios at each point in time. Regressions are performed separately for each of the portfolios, where the portfolio of the most overvalued stocks is Q1, while the most undervalued stocks are in portfolio Q5. Additionally, a spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. Portfolio returns are regressed on the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, LT_Rev, PMU. These factors are obtained from the data libraries from Ken French and Robert Novy-Marx, respectively, and adjusted to be in percent dimensions. The table reports the regression coefficients and associated *t*-statistics of time-series regressions of portfolio returns (in excess of the industry portfolios based on the 38 Fama French industry classifications) on the factors. The table shows the regression coefficients, associated *t*-statistics, as well as the number of observations and R-Squared. *, **, and *** indicate statistical significance at the 10% (5%, 1%) significance level. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 12/1977-12/2012. All variables are defined in Appendix B.

(continued)

Table 7: Buy-and-Hold Returns (continued)

	Q1 (Overvalued)			Q2			Q3			Q4			Q5 (Undervalued)			Q5-Q1 (Undervalued - Overvalued)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Industry-Adjusted Return	-0.1926	[-2.35]	**	-0.1343	[-1.37]		-0.0220	[-0.22]		0.0605	[0.58]		0.1530	[1.67]	*	0.3457	[3.18]	***
Intercept	-0.2348	[-2.91]	***	-0.1284	[-1.42]		-0.0545	[-0.58]		-0.0127	[-0.13]		0.1074	[1.40]		0.3422	[3.61]	***
Mkt_RF	0.0402	[2.12]	**	0.0203	[0.96]		0.0378	[1.71]	*	0.0389	[1.76]	*	0.0001	[0.00]		-0.0401	[-1.80]	*
SMB	-0.1735	[-5.87]	***	-0.2743	[-8.27]	***	-0.2710	[-7.87]	***	-0.2113	[-6.12]	***	-0.0813	[-2.88]	***	0.0922	[2.65]	***
HML	-0.0535	[-1.67]	*	-0.1767	[-4.93]	***	0.0132	[0.35]		0.2461	[6.59]	***	0.3472	[11.39]	***	0.4007	[10.65]	***
Mom	0.0793	[4.45]	***	0.1144	[5.72]	***	0.1147	[5.52]	***	0.0734	[3.52]	***	0.0095	[0.56]		-0.0698	[-3.32]	***
ST_Rev	0.0530	[2.18]	**	0.0314	[1.15]		-0.0174	[-0.62]		-0.0190	[-0.67]		-0.0427	[-1.84]	*	-0.0957	[-3.35]	***
LT_Rev	-0.0033	[-0.09]		0.0296	[0.73]		-0.0355	[-0.84]		-0.1118	[-2.64]	***	-0.1722	[-4.98]	***	-0.1689	[-3.96]	***
RSquare	0.13			0.23			0.22			0.27			0.37			0.32		
Observations	421			421			421			421			421			421		
Intercept	-0.3339	[-4.16]	***	-0.2094	[-2.29]	**	-0.1750	[-1.88]	*	-0.1643	[-1.80]	*	0.0082	[0.11]		0.3421	[3.51]	***
Mkt_RF	0.0493	[2.67]	***	0.0277	[1.32]		0.0489	[2.28]	**	0.0528	[2.52]	**	0.0092	[0.52]		-0.0401	[-1.79]	*
SMB	-0.1548	[-5.37]	***	-0.2590	[-7.88]	***	-0.2483	[-7.41]	***	-0.1828	[-5.56]	***	-0.0626	[-2.28]	**	0.0922	[2.63]	***
HML	-0.0010	[-0.03]		-0.1337	[-3.61]	***	0.0771	[2.04]	**	0.3265	[8.83]	***	0.3998	[12.95]	***	0.4008	[10.14]	***
Mom	0.0791	[4.58]	***	0.1143	[5.81]	***	0.1145	[5.71]	***	0.0732	[3.72]	***	0.0094	[0.57]		-0.0698	[-3.32]	***
ST_Rev	0.0405	[1.72]	*	0.0212	[0.79]		-0.0326	[-1.19]		-0.0381	[-1.42]		-0.0552	[-2.46]	**	-0.0957	[-3.33]	***
LT_Rev	-0.0275	[-0.78]		0.0099	[0.24]		-0.0649	[-1.58]		-0.1488	[-3.69]	***	-0.1964	[-5.84]	***	-0.1689	[-3.92]	***
PMU	0.1823	[5.34]	***	0.1490	[3.83]	***	0.2216	[5.59]	***	0.2789	[7.17]	***	0.1824	[5.62]	***	0.0002	[0.00]	
RSquare	0.19			0.26			0.28			0.35			0.41			0.32		
Observations	421			421			421			421			421			421		

Table 8: Revisions in Accounting Data

The table shows results from factor model time-series regressions. The mispricing signal is created based on alternatively the regular Compustat or the As First Reported (AFR) Compustat, using the 24 of the 28 accounting variables that are also available on the AFR database. Stocks are sorted each month into quintiles based on the mispricing signal and combined into equally-weighted portfolios. Regressions are performed separately for each of the portfolios, where the portfolio of the most overvalued stocks is Q1, while the most undervalued stocks are in portfolio Q5. Additionally, a spread portfolio is formed as the difference between the returns of the portfolios Q5 and Q1. Portfolio returns are regressed alternatively on an intercept (Industry-Adjusted Return), on the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, and LT_Rev (6-Factor Model), or on the excess return on the market portfolio (Mkt_RF), SMB, HML, Mom, ST_Rev, and LT_Rev and PMU (7-Factor Model). These factors are obtained from the data libraries from Ken French and Robert Novy-Marx, respectively, and adjusted to be in percent dimensions. The table reports the regression coefficients and associated *t*-statistics of time-series regressions of portfolio returns (in excess of the industry portfolios based on the 38 Fama French industry classifications) on the factors. The table shows the regression coefficients, associated *t*-statistics, as well as the number of observations and R-Squared. *, **, and *** indicate statistical significance at the 10% (5%, 1%) significance level. The sample consists of all ordinary common stocks of U.S. nonfinancial firms listed on a major exchange (NYSE, Amex, Nasdaq) with a share price at the beginning of the return month of not less than USD 5. The sample period is 1/1988-12/2012. All variables are defined in Appendix B.

	Q1 (Overvalued)			Q2			Q3			Q4			Q5 (Undervalued)			Q5-Q1 (Undervalued - Overvalued)		
	Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat		Coef	*t*-stat	
Regular Compustat																		
Industry-Adjusted Return	-0.4625	[-3.69]	***	-0.1802	[-1.19]		-0.0140	[-0.10]		0.0985	[0.70]		0.3121	[2.41]	**	0.7746	[5.10]	***
6-Factor Alpha	-0.6063	[-5.41]	***	-0.3232	[-2.56]	**	-0.1281	[-1.04]		0.0285	[0.23]		0.2978	[2.71]	***	0.9041	[7.18]	***
7-Factor Alpha	-0.6762	[-5.91]	***	-0.4002	[-3.10]	***	-0.2498	[-2.03]	**	-0.1263	[-1.05]		0.1954	[1.76]	*	0.8716	[6.72]	***
As First Reported Compustat																		
Industry-Adjusted Return	-0.4224	[-3.32]	***	-0.1021	[-0.67]		-0.0427	[-0.30]		0.0500	[0.36]		0.2712	[2.13]	**	0.6935	[4.64]	***
6-Factor Alpha	-0.5748	[-5.12]	***	-0.2319	[-1.81]	*	-0.1501	[-1.26]		-0.0142	[-0.12]		0.2397	[2.18]	**	0.8145	[6.77]	***
7-Factor Alpha	-0.6387	[-5.57]	***	-0.3169	[-2.42]	**	-0.2800	[-2.36]	**	-0.1646	[-1.37]		0.1432	[1.29]		0.7819	[6.31]	***

Appendix A: Discussion and Proof of Result in the Paper's Introduction

For a given date, let X^* denote the NxK matrix of K accounting variables for each of N firms, with $K < N$. The accounting variables are reported (or transformed) at the firm level (*i.e.*, earnings, dividends, depreciation, and book equity for the firm as a whole rather than per share), preserving the linearity of valuation.[1] Thus, the accounting items that would be reported for an investment represented by the $1xN$ vector w in the N firms would be the $1xK$ vector wX^*. More generally, for N distinct investments given by the rows of the NxN matrix W, the accounting statements of the investments would be given by the rows of the NxK matrix WX^*. Thus, the N replicating investments must satisfy $WX^* = X^*$, and if the associated fair value estimates are further required to have average mispricing of zero, then W must satisfy $WX=X$, where the $Nx(K+1)$ matrix X is X^* augmented by a (first) column of 1's.[2] With entries that are functions of X, W's rank deficiency leads to an infinite number of Ws that perfectly replicate each of the N targets' accounting items while producing zero average mispricing.

Proposition: There is a unique W of rank K+1 that is a function of X that produces zero average mispricing and also minimizes the mean-squared prediction error of any non-accounting attribute v of the targets. This is the one given by the idempotent projection matrix statisticians are so familiar from linear regression.

---

[1] For example, the revenue of an investment that buys up 100 per cent of two firms is the sum of their revenues; the earnings of an investment that is 50% of investments A and B is the average of A and B's earnings. Linearity in the portfolio mathematics of accounting items from firm combinations views these combinations as ETFs rather than as full-fledged mergers or acquisitions. Mergers often have synergies, and purchase accounting treatment allocates goodwill to the balance sheet items of the target. Such synergies and accounting treatments generally violate the linearity discussed here.

[2] This means that the N eigenvalues of W consist of K+1 "1"s and N-K-1 "0"s. Moreover, the eigenvectors of W associated with the eigenvalue of 1 consist of the cross-section of each of the K accounting variables and an N-vector of 1's, as well as any linear combination of these eigenvectors. The 1 vector as eigenvector implies W's "weights" sum to one, which is isomorphic to a market portfolio that is never estimated as mispriced.

Proof: Project any variable y not spanned by X onto X, which decomposes $y = X(X^TX)^{-1}X^Ty + \varepsilon$, with the vector ε orthogonal to X. Then, the quadratic minimization problem of finding W with eigenvectors X for eigenvalue 1 that minimizes the sum of squared errors simplifies to choosing the weight matrix W that minimizes $[X(X^TX)^{-1}X^T - W)y + \varepsilon]^T[X(X^TX)^{-1}X^T - W)y + \varepsilon]$, which trivially forces W to be the least squares projection matrix, irrespective of the value of the vector y. Since ε is orthogonal to X and mean zero in sample, it must be orthogonal to W if W is assumed to depend only on X.■

Setting

$$W = X(X^TX)^{-1}X^T$$

predicts a cross-section of the attribute v, denoted P, that is the least squares prediction, *i.e.*,

$$P = Wv = X(X^TX)^{-1}X^Tv.$$

Appendix B: Variable Definitions

The table shows the variable name (or mnemonic), the description (or construction) of the data item, as well as the source (database). CRSP and Compustat are from the merged database on WRDS.

Variable	Definition	Source
ATQ	Assets - Total	Compustat
IBADJQ	Income Before Extraordinary Items - Adjusted for Common Stock Equivalents	Compustat
IBCOMQ	Income Before Extraordinary Items - Available for Common	Compustat
IBQ	Income Before Extraordinary Items	Compustat
LSEQ	Liabilities and Stockholders Equity - Total	Compustat
DVPQ	Dividends - Preferred/Preference	Compustat
NIQ	Net Income (Loss)	Compustat
SEQQ	Stockholders Equity > Parent > Index Fundamental > Quarterly	Compustat
REVTQ	Revenue - Total	Compustat
SALEQ	Sales/Turnover (Net)	Compustat
XIDOQ	Extraordinary Items and Discontinued Operations	Compustat
CSTKEQ	Common Stock Equivalents - Dollar Savings	Compustat
PPENTQ	Property Plant and Equipment - Total (Net)	Compustat
DLTTQ	Long-Term Debt - Total	Compustat
CEQQ	Common/Ordinary Equity - Total	Compustat
PSTKQ	Preferred/Preference Stock (Capital) - Total	Compustat
NOPIQ	Non-Operating Income (Expense) - Total	Compustat
DOQ	Discontinued Operations	Compustat
XIQ	Extraordinary Items	Compustat
LTMIBQ	Liabilities - Total and Noncontrolling Interest	Compustat
LTQ	Liabilities - Total	Compustat
LCTQ	Current Liabilities - Total	Compustat
ACTQ	Current Assets - Total	Compustat
ANCQ	Non-Current Assets - Total	Compustat
PIQ	Pretax Income	Compustat
TXTQ	Income Taxes - Total	Compustat
AOQ	Assets - Other - Total	Compustat
LOQ	Liabilities - Other	Compustat
SharePrice	Stock price (in dollar and cents)	CRSP
Number of Shares Outstanding	Number of shares outstanding (in millions)	CRSP
Return	Monthly Stock Return	CRSP
Beta	Annual Market Beta	CRSP
Industry Classification	38 industries	Ken French website
Industry Portfolios	Monthly returns on 38 industry portfolios	Ken French website
Mkt_RF	Monthly market index return net of risk-free rate	Ken French website
SMB	Monthly Small Minus Big (SMB) portfolio return	Ken French website
HML	Monthly High Minus Low (HML) portfolio return	Ken French website
Mom	Monthly Momentum portfolio return	Ken French website
ST_Rev	Monthly Short-term reversal portfolio return	Ken French website
LT_Rev	Monthly long-term reversal portfolio return	Ken French website
PMU	Monthly profitability factor	Novy-Marx website

(continued)

Appendix B: Variable Definitions (continued)

Variable	Definition
SUE	Quarterly earnings surprise based on a rolling seasonal random walk model (Livnat and Mendenhall, 2006, page 185)
Accruals	Accruals = [NOA(t)-NOA(t-1)]/NOA(t-1), where NOA(t) = Operating Assets (t) - Operating Liabilities (t). Operating Assets is calculated as total assets (ATQ) less cash and short-term investments (CHEQ). Operating liabilities is calculated as total assets (ATQ) less total debt (DLCQ and DLTTQ) less book value of total common and preferred equity (CEQQ and PSTKQ) less minority interest (MIBTQ) (Richardson et al., 2001, p. 22)
Gross Profitability	(Revenue(REVTQ) - Cost of Goods Sold(COGSQ))/Total Assets(ATQ) (Novy-Marx 2013)
Default Risk	Default probability from Merton (1974) model
Market Capitalization	Stock Market Capitalization of Common Stock, calculated as product of Share Price(PRC) * Number of Shares Outstanding(SHROUT)
Book/Market	(Book Equity(CEQQ) + Deferred Taxes Balance Sheet(TXDBQ))/MarketCapitalization
Mispricing Percentage (M)	-1 * Residual/ MarketCapitalization