Journal of Financial Economics 128 (2018) 125–147

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Agnostic fundamental analysis works[☆]

Söhnke M. Bartram[a,*], Mark Grinblatt[b]

[a] Warwick Business School, Finance Group, University of Warwick, Coventry CV4 7AL, UK
[b] UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095-1481, USA

A R T I C L E I N F O

Article history:
Received 10 June 2016
Revised 17 October 2016
Accepted 25 November 2016
Available online 15 December 2017

JEL classification:
G11
G12
G14

Keywords:
Valuation
Asset pricing
Market efficiency
Fundamental analysis
Point-in-Time
Theil-Sen

A B S T R A C T

To assess stock market informational efficiency with minimal data snooping, we take the view of a statistician with little knowledge of finance. The statistician uses techniques such as least squares to estimate peer-implied fair values from the market values of replicating portfolios with the same accounting statements as the company being valued. Divergence of a company's peer-implied value estimate from its market value represents mispricing, motivating a convergence trade that earns risk-adjusted returns of up to 10% per year and is economically significant for both large and small cap firms. The rate of convergence decays to zero over the subsequent 34 months.

☆ We would like to thank William Schwert (the editor), and Richard Sloan (the referee) as well as Tom Aabo, Tim Adam, David Aboody, Senay Agca, Hank Bessembinder, Michael Brennan, Greg Brown, Adrian Buss, Ian Cooper, Tarun Chordia, Jean Dermine, Bernard Dumas, Alex Edmans, Peter Feldhütter, Wayne Ferson, Richard Frankel, Federico Gavazzoni, Michael Goldstein, Robin Greenwood, Denis Gromb, John Hall, Luigi Guiso, Larry Harris, David Hirshleifer, Christopher Hrdlicka, Samuli Knupfer, Ralph Koijen, Sonia Konstantinidi, Alok Kumar, John Kuong, Wei-Hsien Li, Juhani Linnainmaa, Karl Lins, Marc Martos-Vila, James Ohlson, Anna Pavlova, Neil Pearson, Fulvio Pegoraro, Jeffrey Pontiff, Scott Richardson, Giuliano De Rossi, Rik Sen, Siew Hong Teoh, Sheridan Titman, Raman Uppal, Mathijs van Dijk, Theo Vermaelen, Russ Wermers and seminar participants at Aarhus University, Barclays Capital, Cardiff Business School, Chinese University of Hong Kong, Einaudi Institute for Economics and Finance (EIEF), ESCP Europe, European School of Management and Technology, Frankfurt School of Finance and Management, George Washington University, Goethe University Frankfurt, Hong Kong University of Science and Technology, European Institute of Business Administration (INSEAD), Kühne Logistics University, London Business School, Manchester Business School, National Central University Taiwan, Norges Bank Investment Management, Princeton University, Queen Mary University of London, University of Illinois at Chicago, University of New South Wales, University of Warwick, University of Washington, Centre for Economic Policy Research (CEPR) First Annual Spring Symposium in Financial Economics, Ninth Financial Risks International Forum on New Challenges Facing the Investment Management Industry, the 2017 Western Finance Association Annual Conference, 2016 annual meetings of the American Economic Association, 2015 Axioma Quant Forum, 2015 Inquire UK seminar, 2015 LA Finance Day, 2014 GEA conference in Kiel, Germany, the Ph.D. Nordic Finance Workshop in Stockholm, Sweden, and the Seventh Conference on Professional Asset Management in Rotterdam, the Netherlands, for helpful comments and suggestions. Bartram gratefully acknowledges the warm hospitality of the UCLA Anderson School of Management, NYU Stern School of Business, London Business School, Center for Financial Studies (CFS)/House of Finance at Goethe University, EIEF and European University Institute during visits to these institutions. The authors thank James A. Ohlson and Seil Kim for helpful discussion of and providing Statistical Analysis System (SAS) code for the Theil–Sen estimation. Shao (Ken) Chang and Patrick Kiefer provided excellent research assistance. Kevin Aretz kindly provided data on default probabilities.

* Corresponding author.
E-mail address: s.m.bartram@wbs.ac.uk (S. Bartram).

https://doi.org/10.1016/j.jfineco.2016.11.008
0304-405X/

1. Introduction

A cornerstone of market efficiency is the principle that trading strategies derived from public information should not work. 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, profitability, book-to-market, and a host of other signals have earned risk-adjusted profits in the past.[1] However, the motivation for studying these signals is not always apparent.[2]

The motivation for studying whether fundamental analysis works is more obvious. 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 indicating deviations from fair value. Abnormal profits arise from convergence to fair value, at one extreme via short-term price movements toward fair value or, more slowly, via distributions of dividends, takeovers, private buyouts, or asset liquidation. 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.

Most studies of fundamental analysis require highly stylized models of fair value. For example, discounted cash flow models of fundamental value require near-term forecasts of cash flows, short-term and long-term cash flow growth rates, and proper discount rates. The researcher's broad discretion over the choice of model and its parameters makes it difficult to assess whether fundamental analysis works or if the observed abnormal profits are the outcome of a data snooping exercise.[3] The researcher can plausibly argue that the deviations of market prices from fair value underlie the predictability of future returns only with a more agnostic and less discretionary approach to fundamentals-based equity valuation.

Despite of the popularity of stylized models to assess fair value, 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 models and parameter estimates. Our simple and agnostic approach to fundamental analysis approximates a company's fair equity value as a linear function of virtually all of its most recently reported balance sheet, income statement, and cash flow statement items. Our only restriction is that the function's coefficients, which are determined each month and the same across firms, offer the lowest degree of mispricing (as measured by variance) of a randomly selected investment dollar in the economy that month. The term "peer-implied fair value" seems appropriate here as each firm's fair value prediction is based on contemporaneous valuations of all firms in the stock market.

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. Each month's peer-implied fair values are the predictions of monthly cross-sectional regressions of market capitalizations on all firm-level accounting items with broad coverage across firms. The accounting items are known to market participants at the time according to the Compustat Point-in-Time (PIT) data set, which we exclusively use throughout the study. Regression residuals identify which firms' market capitalizations deviate from month-specific valuation norms that are attached to a firm's accounting statement. The market portfolio is considered to be fairly valued at all times because regression estimates with a constant force the sum of the residuals to zero.

This alternative approach to fundamental analysis is unorthodox, but it helps avoid the temptation to data snoop across model specifications and parameter estimates. The theorist's license in more traditional fundamental analysis is best suspended when it offers too much discretion over implementation, potentially yielding significant results by chance. Predicting fair value as a statistics-constrained linear function of all widely reported accounting items prevents discretion in the selection or weighting of accounting items that could relate to future returns.

After identifying peer-implied values from linear functions of accounting items, we study the profitability of buying undervalued and selling overvalued securities, measured from the percentage deviation of a stock's peer-implied value estimate from its actual market capitalization. Cross-sectional regressions of returns on firm attributes and dummy variables for quintile ranks, as well as time series regressions of mispricing-sorted portfolio returns on factors, assess whether this agnostic take on fundamental analysis offers a profitable investment strategy. We control for industry returns, beta, book-to-market ratios, momentum, short- and long-term reversals, firm size, gross profitability, accruals, earnings surprises, earnings yield, and 12 other known anomalies, including alternative mispricing estimates. The abnormal return (alpha) spreads between portfolios of the extreme quintiles of stocks sorted on percentage misvaluation are between 4% and 10% per year, depending on the risk adjustment

[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 such as 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, and return covariation within the value and growth categories, embodied in the HML (high minus low) 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).

[3] The literature notes that return predictability and contemporaneous correlations between returns and risk-adjustment factors could be the spurious outcome of specification search, data snooping, sample selection bias, and repeated hypothesis testing. See, for example, Kogan and Tian (2013), McLean and Pontiff (2016), Schulmeister (2009), Sullivan, Timmerman, and White (1999), Foster, Smith, and Whaley (1997), Fama (1991), Lo and MacKinlay (1990), Heckman (1979), and Leamer (1978). Studies that examine return predictors in out-of-sample periods do not necessarily resolve the data snooping dilemma, as the specific anomalies studied are highly correlated with other anomalies known to work in the same out-of-sample periods.

procedure used, positive in almost 60% of the 310 months studied, and prevalent in large and small firms (although value weighting reduces risk-adjusted alphas by about one-third). The mispricing regressor is of similar significance as momentum, and it surpasses the greater significance hurdles suggested by Harvey et al. (2016) and Green et al. (2013).[4]

Despite the handicaps imposed, our statistician's approach to valuation 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 values obtained with our approach are the market values of synthetic stocks or replicating portfolios; replicating, because each of the portfolios' fundamental characteristics is identical to those of the firm being valued.[5] Because the number of firms N is large relative to the rank K of an N x K 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 value prediction and the replicating portfolio matrix attached to it can be deduced from three appealing assumptions:

(1) The $N \times N$ replicating portfolio matrix has weights on stocks that make the average valuation error zero (forcing the market portfolio to appear fairly priced in relative terms each month).
(2) The replicating portfolio matrix has weights that are functions only of the K-dimensional accounting information and are not functions of firms' market capitalizations, returns, or other variables besides the accounting information.
(3) The replicating portfolio matrix minimizes the average squared deviation across securities of any attribute (including market capitalization) not spanned by the K-dimensional accounting attributes.

The set of replicating portfolio weights satisfying the above criteria forms an $N \times N$ idempotent projection matrix $X(X^TX)^{-1}X^T$, tied to the cross-sectional regression of market capitalization on X. 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, instead of fitting market capitalization, it was designed to fit earnings growth rates, the age of the chief executive officer, the latitude of the firm's headquarters, or the returns of a stock (see, for example, Bessembinder et al., 2015). Despite market capitalization's nonexistent role in the replicating portfolio, the market values of the replicating portfolios capture all of the dynamics of the relation between the actual market values of firms and their accounting variables. And, in contrast to prior studies that predict returns from specific variables of interest, such as 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.

Convergence of market prices to their peer-implied fair value is the most likely source of the profitability we uncover from this trading strategy. Admittedly, higher discount rates imply low market values and vice versa, other things equal.[6] The mechanical relation in the cross section between market values (or ratios involving market values such as 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 relation does not identify whether the observed return spreads, tied to this type of anomaly, are due to differences in risk or pricing errors. We present evidence suggesting that convergence to fair value, not risk differences, accounts for the efficacy of our mispricing signal. The mispricing signal has no strong relationships with known risk factors, the signal ranks decay more quickly than risk attributes, and the signal efficacy decays to zero over the subsequent 34 months.[7]

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, income statement, and cash flow statement items reported by sufficient numbers of firms. The large numbers of highly (or perfectly) collinear variables implies that coefficient signs 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 peer-implied fair value approach used here is unlikely to be a superior mousetrap for capturing the intrinsic values of securities. However, if the crude statistician's approach to fundamental analysis works, then more accurate ways of measuring mispricing should work even better.

With this in mind, we apply the same model but also estimate fair values with an alternative to least squares on

[4] In Harvey, Liu, and Zhu (2016), 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.

[5] As Ross (1978, p. 455) notes, even the simplest discounting of risk-free cash flows is a comparison between the traded price of a quantity of risk-free bonds available in securities markets and an asset that produces a future risk-free cash flow. In the capital asset pricing model (CAPM), a stock's 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 probability-weighted pricing kernels used to obtain fair values of all assets. And even when parameters such as risk aversion are estimated from experiments, the lotteries used to obtain those parameters are deemed to be fairly valued.

[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).

[7] Our approach has similarities to pairs trading strategies, as popularized by Gatev, Goetzmann, and Rouwenhorst (2006). However, in contrast to pairs trading, our strategy does not invest in the replicating portfolio, which is used only as a peer-implied valuation benchmark for each stock.

all firms that is more robust to outliers. This alternative is inspired by the estimator developed by Theil (1950) and Sen (1968) (TS) that is based on median coefficients across a large number of perfectly fit slopes from subsets of the sample. Given the high correlations between the regressors, the coefficients and thus their median are not uniquely identified in our setting. We address this issue by adapting the TS methodology to the median fair value for each firm, which is uniquely identified, as opposed to using the median coefficients to derive the predictions. As a more robust and presumably better estimation technique, a firm's median fair value prediction should give more accurate predictions of its true fair value in each month. The paper's TS-inspired estimates of fair value generally lead to higher trading profits than least squares estimates derived from one regression fit to all firms that month (the ordinary least squares (OLS) prediction). However, both estimators of fair value generate significantly positive alphas.[8]

The paper is organized as follows: Section 2 discusses the related literature. Section 3 describes the data and the empirical methodology used to estimate mispricing. Section 4 studies Fama-MacBeth regressions of return predictability based on the mispricing signal, and computes signal-sorted alphas from factor models. Section 5 offers evidence that convergence to fair value, as opposed to risk, is the source of the abnormal returns earned from our approach to fundamental analysis. Section 6 investigates how the mispricing signal relates to a set of 22 documented anomalies and provides direct analysis of the alpha-generating role played by each of the 28 accounting items. Section 7 concludes the paper.

2. Related literature

Direct study of whether the estimation of fair market values per se leads to trading strategies that can earn abnormal profits is rare. Bhojraj and Lee (2002), Liu et al. (2002), and Cooper and Lambertides (2014) examine the relative valuation of target and comparable firms in the context of multiples valuation. However, Cooper and Lambertides (2014) find no evidence of predictability, and the other two papers do not investigate whether mispricing can be used to generate profitable trading strategies.

Notable exceptions to the dearth of direct research on whether fundamental analysis works are studies that estimate fair value from the residual income model, including Ohlson (1990, 1991, 1995), Dechow et al. (1999), and particularly Frankel and Lee (1998) and Lee et al. (1999). Frankel and Lee (1998) show that deviations from fair value predict long-term returns, especially between 24 and 36 months after receiving the mispricing signal. Lee et al. (1999) show that ratios of fair value to price predict the Dow 30 index. These papers differ from ours in their use and discretionary implementation of a specific theoretical model of intrinsic value (which requires analyst forecasts, aggregation and extrapolations of the forecasts, and discount rates). We do not require discretionary decisions about how to aggregate analyst forecasts or about how income growth rates and discount rates vary over time, forecast horizons, and firms. In this sense, our statistician's approach to fair value estimation can be theoretically consistent with the residual income model, but it is conceptually distinct from the implementation of these models in the literature. Moreover, our mispricing signal retains predictive power after controlling for these alternative estimates of mispricing.[9]

A large number of studies predict stock returns using information from financial statements (e.g., the ratio of investment to sales), often combined into a ratio with the firm's stock price (e.g., dividend yield). In fact, research on anomalies has identified more than 333 known predictors of future returns, of which many are based on items from accounting statements.[10] The closest to a study of fundamental analysis using a large number of accounting variables is Ou and Penman (1989), which considers accounting variables as predictors of future earnings changes and shows that the logit-estimated probability of an earnings increase forecasts stock returns. Holthausen and Larcker (1992) show that the Ou-Penman methodology can be improved by weighting the same accounting variables to best predict returns (but not earnings) out-of-sample.

In another example, Abarbanell and Bushee (1998) study the April to March returns of firms with December fiscal year-ends and find that the weighted averages of the ranks for changes in nine accounting variables predict a firm's return.[11] They select discretionary accounting constructs and weights solely because they predict returns in sample, not because the accounting constructs offer an estimate of a firm's fair value. While these studies are like ours in that they predict returns from accounting data, their signals are not based on deviations of intrinsic value estimates from market prices. Even with Ou and Penman (1989), no *a priori* reason exists to think that a high probability of an earnings increase would lead to a high return unless the market was forecasting a lower earnings increase probability, and there is no direct evidence of the market's earnings forecast in the Ou and Penman (1989) paper. The only indirect evidence that it is lower (that is, the high return achieved) does not survive

[8] We also consider valuation regressions without an intercept. While these models no longer ensure that the market portfolio is fairly priced at each point in time, they do have the appealing feature that firms with zero values for all accounting items have a valuation of zero. Risk-adjusted quintile spreads for the resulting misvaluation signals are essentially the same (TS estimate of fair value) or somewhat to moderately stronger (OLS estimate of fair value) without a constant than with a constant.

[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 US stocks. Deviations from fair value have also been used to study misvaluation and Q theories of mergers and acquisition activity and agency costs, based on residual income valuation (e.g., Dong, Hirshleifer, Richardson, and Teoh, 2006) or (annual, industry-level) cross-sectional regressions of valuation metrics (Tobin's q, market capitalization) on determinants of fundamental value (e.g., Edmans, Goldstein, and Jiang, 2012; Rhodes-Kropf, Robinson, and Viswanathan, 2005; Habib and Ljungqvist, 2005).

[10] See, for example, Green, Hand, and Zhang (2013), Harvey, Liu, and Zhu (2016), and McLean and Pontiff (2016).

[11] Related papers include Greig (1992), Holthausen and Larcker (1992), Lev and Thiagarajan (1993), Abarbanell and Bushee (1997), Piotroski (2000), Mohanram (2005), and Piotroski and So (2012).

beyond the Ou-Penman sample period (as Holthausen and Larcker (1992) show).

The large number of metrics and justifications across studies makes it difficult to determine which approach is correct or robust. The selection of return predictors, proper mix for a trading signal, and success at publication hinges on the ability of the signal to predict returns. Like these other papers, we relate our signal to future returns, but the motivation for our hypothesis is uncomplicated and transparent: Deviations from fair value are more likely to contract than expand.

Finally, while many fund managers use fundamental analysis, we cannot infer whether fundamental analysis works by measuring the performance of professional money managers. Their alphas, which are on the order of 0–100 basis points per year before deducting transaction costs, fees, and other expenses, are small.[12] More important, we do not know which managers use fundamental analysis to buy stocks their models claim are undervalued, and which employ strategies that are completely different.

3. 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. We compute each stock's degree of under- or overvaluation for the last day of every month in the sample. We then track the returns of stocks over the subsequent month[13] and relate these returns to the stock's beginning-of-month mispricing.

3.1. Sample period and data filters

Our sample contains 310 return months (March 1987 through December 2012) and, thus, 310 portfolio formation dates starting Friday, February 27, 1987 and ending Friday, November 30, 2012.[14] On the day of mispricing measurement and portfolio formation the stock must be in the Center for Research in Securities Prices (CRSP) Monthly Stock File as the only common equity share class of a US corporation (share classes 10 and 11) and be listed on the New York Stock Exchange, American Stock Exchange, or Nasdaq Stock Market–National Market System (NASDAQ–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); have positive total assets and all accounting inputs required for the trading signal publicly disclosed at the portfolio formation date; and possess a Standard Industrial Classification (SIC) code that is not financial services (codes 60–69).

3.2. Estimating peer-implied fair value and mispricing

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 portfolio formation dates t and each stock j, we calculate a mispricing signal,

$$M_{j,t} = \frac{P_{j,t} - V_{j,t}}{V_{j,t}} \tag{1}$$

as 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 peer-implied values inferred from their most recent accounting statements. Such stocks are expected to outperform the overpriced stocks in the future. Stocks with highly negative $M_{j,t}$ are overvalued stocks that are expected to under-perform. By construction, the date t market capitalization-weighted average of $M_{j,t}$ is zero.

To economically quantify the effect of mispricing, we rank each of the regression's stocks at the end 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 dummy variables can be interpreted as the added return from belonging to the respective mispricing quintile compared with the Q1 quintile.

The regressors for the date t fair value predictions come from stock j's (and other firms') most recently reported 10-Q or 10-K balance sheet, income or cash flow statement items, obtained from the CRSP-merged Compustat Point-in-Time database. This database captures data values known by the market from both preliminary announcements and final sources and combines originally reported and restated data in annual, quarterly, and year-to-date periodicities. On each of the monthly point dates, including the first, data on up to 20 prior fiscal quarters are available, reflecting the accounting information known to investors.

As is customary when analyzing accounting data, all variables that inform trading positions are winsorized. Here, they are winzorized based on their 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 and including month t. Our results are not sensitive to winsorization.

We employ the 28 most common numerical firm-level accounting items reported by Compustat throughout the sample period listed as coming from the balance sheet (16 items), income statements (11 items), and cash flow statements (1 item) available with four past fiscal quarters on

[12] See, for example, Grinblatt and Titman (1989, 1992, 1993, 1994), Daniel, Grinblatt, Titman, and Wermers (1997), Chen, Jegadeesh, and Wermers (2000), Wermers (2000), Fama and French (2010), Berk and van Binsbergen (2015), and Grinblatt, Jostova, Petrasek, and Philipov (2017). 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 overweighted (compared with the index weight) by actively managed funds greatly outperform those heavily under-weighted after adjustment for risk.

[13] 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 Center for Research in Security Prices (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.

[14] The first point date on the Compustat Point-in-Time database is February 27, 1987. An earlier version of the paper finds similar results using accounting data from the regular, quarterly Compustat going back to 1977. Bartram and Grinblatt (2017) use the mispricing measure to study market efficiency in international equity markets.

February 27, 1987.[15] Appendix B lists these 28 items along with details on variables used in the paper as return regression controls. Twenty-eight items is the largest number of items that achieve a sample of 2000 firms at the sample period's start, which is desirable for statistical precision. This coverage-imposed reduction of the accounting data matrix 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. Five of the 28 accounting items are perfectly spanned by the remaining 23 and, for subsets of firms, a few more are perfectly spanned by the remainder (suggesting that rounding conventions are inconsistent across items and firms). About 92% of the variation in half of the items is captured by the remaining half of the items. While such multicollinearity means the regression coefficients on many of the items are imprecise and sometimes indeterminate, the peer-implied fair value predictions from the 28 items are unique.[16]

We use, at the start of each trading month, the most recently reported 10-K and the most recently reported 10-Qs to identify values for these 28 items. The 16 balance sheet items are from the most recently released accounting statement (10-K or 10-Q). Those from the income and cash flow statements are sums of the quarterly values from the three most recently released 10-Qs and 10-K.[17] Although summing four quarterly values can characterize 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 16 items in the most recent balance sheet and the sum of the four quarterly values of the 12 items derived from the four most recent income and cash flow statements.

Each firm's peer-implied fair value evolves month to month for two reasons. First, market capitalizations, the cross-sectional regression's dependent variable, change, influencing 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 research and development (R&D) become relatively more valuable than the historical norm, as in the 1998–1999 Internet bubble, our cross-sectional approach captures that change in market tastes. Second, firms can report new accounting information during the month. The new information changes the values and coefficients of the regressors used to predict fair value in the next month.

In sum, our peer-based approach to fair value takes no stance 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. Furthermore, it does not rely on a formal theoretical model of fundamental value. Instead, we compare firms with 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.

The OLS fair value regressions have R-squareds that vary month to month. The minimum R-squared (unadjusted for degrees of freedom) is 76.6% (April 2000), the median is 92.7%, and the average is 91.7%. These R-squareds are unimpressive in light of the fact that market capitalization is on the left-hand side and that the right-hand side accounting entries tend to scale with firm size. All estimates of fair value and mispricing are highly inexact, including ours.[18] The noise in our approach implies that our mispricing estimates are larger than the profits from trading on these estimates and that these profits are likely to be enhanced by better mispricing estimation. Also, because sampling error is based on persistent variables, mispricing estimates converge to zero more slowly than true mispricing.

Due to the crude nature of our approach and the unknown statistical properties of the resulting mispricing signal, the empirical analysis focuses on a firm's mispricing quintile. Outliers have two effects here: They distort the distribution of the mispricing variable, particularly in the tails, and they can distort the peer-implied regression coefficients. The first issue is addressed by focusing our analysis on quintile ranks. Alternative estimators to ordinary least squares can also mitigate the impact of outliers on the peer-implied regression coefficients. We employ a variant of the estimator developed by Theil (1950) and Sen (1968) that is robust to outliers.[19] Results for mispricing signals from TS estimation are generally stronger, both economically and statistically, compared with those using mispricing signals from OLS estimation.

In simple estimation settings, the TS estimator is the median coefficient vector from all pairs of prediction vectors and outcomes of the dependent variable for sample subsets that offer a perfect fit. With large numbers of firms and estimated parameters, computational complexity requires estimation by random sampling. Moreover, we focus on the predictions of the regressions as opposed to the regression coefficients to avoid distortions in TS estimates induced by high degrees of multicollinearity. Each draw identifies one hundred random firms in a month and fits the 28 accounting variables plus a constant to the market capitalizations of the firms in that month. For each month, we repeat the draw of one hundred firms 100,000 times before identifying the median of the predicted

[15] Many of the 159 items in Compustat PIT, such as firm name, ticker, and notes, are not numerical. Many are titled "per share."

[16] Alternative mispricing signals that include the prior fiscal quarter of all 28 items help capture growth rates. The performance of these signals is similar if not marginally stronger compared with signals without such lags.

[17] Cash flow items in Compustat PIT are cumulative. We infer the quarterly data item by taking the difference between the cumulative cash flow variables of adjacent fiscal quarters (for fiscal quarters other than the first).

[18] Numeric Investors, an institutional asset manager, estimates fair value from daily regressions of stock prices on an undisclosed set of company fundamentals, analysts' forecasts, etc. (see Perold and Tierney, 1997).

[19] Ohlson and Kim (2015) argue that Theil-Sen estimation is superior to OLS estimation for linear valuation models because it is robust to outliers and does not require variable scaling to improve estimation efficiency. We are grateful to the referee for recommending this approach.

Table 1
Summary statistics by mispricing signal quintiles.
The table reports averages of a number of characteristics of portfolios and firms, including the time series average of the monthly mean characteristics across all firms ("All"), the average monthly cross-sectional correlation of the characteristic with the mispricing signal *M* ("Correlation"), and the average of the monthly 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 US 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 \$5. The sample period is March 1987 to December 2012. All variables are defined in Appendix B.

			Signal quintiles				
Characteristics	All (1)	Correlation (2)	Q1 (overvalued) (3)	Q2 (4)	Q3 (5)	Q4 (6)	Q5 (undervalued) (7)
Mispricing signal (M)	0.9640	1.000	−2.0253	−0.2427	0.3996	1.3042	5.3848
Market capitalization	2,847.7	−0.068	3,541.8	5,941.9	3,006.4	1,365.7	381.3
Book/market	0.5774	0.291	0.4071	0.4198	0.5054	0.6361	0.9186
Beta	0.9280	−0.139	1.0259	1.0018	0.9764	0.9102	0.7227
Accruals	0.9507	−0.010	1.8169	0.8555	0.6942	0.6995	0.7328
SUE	0.0140	0.026	0.0275	0.0021	0.0093	0.0230	0.0082
Gross profitability	0.3685	0.042	0.3011	0.3811	0.3907	0.3875	0.3819
Earnings yield	0.0210	0.159	−0.0537	0.0271	0.0387	0.0439	0.0488
Return from prior month t	2.0692	−0.029	3.5124	2.7653	1.9508	1.3282	0.7908
Return from month $t-1$ to $t-11$	23.41	−0.068	38.741	32.365	21.825	14.64	9.670
Return from month $t-12$ to $t-59$	99.43	−0.048	109.19	114.73	107.47	92.44	73.45

market capitalizations for each firm as the TS estimate of its peer-implied value.

3.3. Summary statistics for the overall sample

Table 1 reports summary statistics describing the relation of the (OLS) mispricing variable *M* to firm size, beta, book-to-market ratio, past returns over a variety of horizons, earnings surprises, accruals, gross profitability, and earnings yield. It reports the time series average of the cross-sectional means of these variables in column (1), the time series average of the correlation of the variable with *M* in column (2), and the time series averages of the means of the variables within five mispricing quintiles. Quintile 1, in column (3), represents the most overpriced stocks, which have average overpricing of 203%. Quintile 5 (column (7)) represents the most underpriced stocks, which have average underpricing of 538%. These figures are large because extreme conditional means sort on sampling error and are therefore biased. However, because our later analysis involves only ranks, no need exists to adjust for the bias with statistical corrections such as Bayesian shrinkage.

As can be seen from Table 1, mispricing is highly related to a number of attributes known to predict returns. Compared with the 20% most overpriced firms, the 20% most underpriced firms are about nine times smaller and have a lower beta, lower past returns (at all three horizons), and higher earnings-to-price and book-to-market ratios. 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.

M has positive correlations with book-to-market and earnings-to-price and negative correlations with firm size and past returns. The negative correlation between beta and estimated mispricing indicates that beta risk could not explain any ability of *M* to forecast average returns (though it would be consistent with Frazzini and Pedersen (2014)). Moreover, while *M's* correlations are below 0.05 with other prominent anomalies in the finance literature, stocks in the most overpriced quintile have greater accruals and (positive) earnings surprises and lower gross profitability than the most underpriced stocks. For this reason, at various points in the paper, we control for the effect of book-to-market, earnings-to-price, size, past returns, gross profitability, accruals, and earnings surprises on future average returns.

4. The mispricing attribute and the cross section of expected returns

This section shows that *M* forecasts returns even after controlling for industry returns, beta, book-to-market ratios, momentum, short- and long-term reversals, firm size, gross profitability, accruals, earnings surprises, and earnings yield.

4.1. Raw returns

Table 2, similar in format to Table 1, addresses the mispricing signal's ability to forecast next-month's return. As indicated in the column headings, the table focuses on mispricing signals based on OLS regressions, but it also reports statistics for signals based on TS regressions. The table reports time series averages of both equal- and value-weighted portfolio returns, the average correlation between the return and mispricing signal, and 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 two similar length subperiods. Table 2 uses the null of efficient markets to test whether the mean return of the most underpriced quintile of stocks (Q5) equals that of the most overpriced quintile. The average difference and associated t-statistic (from the time series of paired differences) appear in column (9) under the OLS heading, flanked on the left in column (8) by the fraction of return differences that are positive.

Table 2
Stock returns and mispricing signal quintiles.

The table reports averages and selected test statistics of monthly portfolio returns, including 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"), and 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 spread between the returns of the most undervalued (Q5) and the most overvalued (Q1) firms, as well as the associated *t*-statistics. 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 against a 50% positive null hypothesis. Columns under the ordinary least squares ("OLS") heading report results for signals from OLS regressions, and columns under the Theil (1950) and Sen (1968) ("TS") heading show results for signals from Theil-Sen regressions as described in the text. Panels A and B report results for equal- and value-weighted portfolios, respectively. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1987 to December 2012.

		OLS								TS
			Signal quintiles					Q5–Q1 (undervalued–overvalued)		Q5–Q1
	All	Correlation	Q1 (overvalued)	Q2	Q3	Q4	Q5 (undervalued)	Fraction > 0 [*p*-value]	Average [*t*-statistic]	Average [*t*-statistic]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Equally weighted portfolios										
Return in month $t+1$	0.9224	0.0050	0.6309	0.9166	0.9713	1.0420	1.0502	58.1	0.4192	0.6790
								[0.00]	[2.38]	[4.03]
Return in month $t+1$ (1987–1999)	1.0575	0.0071	0.7122	1.1522	1.1609	1.1850	1.0763	58.4	0.3641	0.5622
								[0.04]	[1.56]	[2.41]
Return in month $t+1$ (2000–2012)	0.7889	0.0030	0.5507	0.6840	0.7841	0.9009	1.0244	57.7	0.4737	0.7943
								[0.05]	[1.79]	[3.26]
Panel B: Value-weighted portfolios										
Return in month $t+1$	0.8669	0.0091	0.4753	0.8964	0.8971	1.0519	1.0217	54.5	0.5465	0.4850
								[0.11]	[2.53]	[1.67]
Return in month $t+1$ (1987–1999)	1.4113	0.0013	1.0979	1.4080	1.4248	1.3505	1.2417	52.6	0.1437	−0.0474
								[0.52]	[0.57]	[−0.13]
Return in month $t+1$ (2000–2012)	0.3294	0.0168	−0.1394	0.3914	0.3763	0.7571	0.8046	56.4	0.9440	1.0106
								[0.11]	[2.72]	[2.24]

For comparison, column (10) under the heading TS reports quintile spreads and associated t-statistics for signals from TS regressions.

The average correlation between a firm's signal and its future returns is 0.0050. Moreover, average returns are nearly perfectly monotonic in the mispricing quintiles, both in the full sample period and for subperiods, using both equal- and value-weighted portfolios. The next-month return spread between the least and most underpriced stock quintiles is 0.42% (0.55% when value weighted), an annualized return spread of 5.0% per year (6.6% per year for the spread in the value-weighted portfolios). Finally, the Q5–Q1 spread is positive in about 58% of the months (55% when value weighted). Across subperiods, return spreads are largest in the years 2000–2012. Using the Theil-Sen approach to measure peer-implied value mildly enhances the mispricing signal's efficacy in several cases compared with OLS signals, yielding spreads between the most under- and overvalued stocks of 8.2% and 5.8% per year for equally weighted and value-weighted portfolios for the full sample period, 12.1 % and 9.5% in the last 12 years of the sample.

4.2. Fama-MacBeth cross-sectional regressions

The Table 2 raw return differences could 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 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

$$R_{j,t+1} = a_t + b_t M_{j,t} + \sum_{s=1}^{S} c_{s,t} X_{j,s,t} + e_{j,t+1}, \tag{2}$$

where $R_{j,t+1}$ is stock j's month $t+1$ return and $X_{j,s,t}$ is the end-of-month t value of firm j's control characteristic s including industry fixed effects.[20]

Table 3's time series averages of coefficients have Fama and MacBeth (1973) t-statistics, which appear in brackets. Because we do not know the correct functional form for the mispricing signal and want to calibrate its economic effect, the regressions use quintile dummies (Q2, Q3, Q4, and Q5, with Q1 omitted due to the regression intercept) for all of the anomalies instead of the variables themselves. For brevity, the table displays coefficients and test statistics only for the Q5 dummy, which represents the difference in returns from being in Q5 compared with Q1. Dummy coefficients for Q2, Q3, and Q4 for each of the characteristics are also included in the regression but not reported.

Table 3 reports several specifications to assess the mispricing signal's ability to predict returns. Its first specification lacks controls for other characteristics besides industry. The second and third add the more classic characteristics of beta, size, book-to-market, and past returns (over three nonoverlapping horizons). The fourth and fifth add the four characteristics gross profitability, accruals, earnings yield, and earnings surprises. These five specifications of Table 3 are based on OLS signals, and the sixth specification has the same control variables as Specification 5 but uses Theil-Sen signals. To facilitate comparisons across specifications, month t's regressions omit firms lacking data for all specifications. Results are highly similar without this restriction.

Specification 1 indicates that the average industry-adjusted return of mispricing Quintile 5 exceeds that of Quintile 1 by 46 basis points per month. M's Q5–Q1 monthly spread increases to 54 basis points with traditional controls (Specification 3), which is five times the book-to-market effect in the same specification. In the kitchen sink regression (Specification 5), the characteristics controlled quintile spread for the mispricing signal is 36 basis points when it is based on OLS regressions. For the Theil-Sen signal, the effect is 44 basis points, about four–fifth the size of the same specification's momentum effect (57 basis points). In all specifications, mispricing significantly predicts next month's return. Mispricing's smallest coefficients of 0.362 and 0.435 appear in the kitchen sink Specifications 5 and 6. The coefficient is larger and more significant ($t = 4.37$) with the traditional controls of Specification 3. The coefficients are similar if we do not adjust for industry effects.

Clearly, Specifications 5 and 6's combination of book-to-market with other less traditional controls tied to earning captures some of the return predictability attributed to mispricing (and momentum). The results here indicate that the mispricing signal generates significant characteristics-adjusted profits that are about two-thirds of Specification 3's profits, which lacks the earnings-related controls. The fact that the accounting variables and their weighting are chosen with no input from future returns suggests that less ad hoc approaches to fundamental analysis are likely to prove even more fruitful than what we propose.[21]

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 B of Table 3 averages Panel A coefficients for the subperiod 1993–2012. In the four specifications that employ our mispricing signal, the effect of mispricing in the subperiod are about the same (TS) or stronger (OLS) than in the full sample period. By contrast, Panel B shows mod-

[20] 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.

[21] Moreover, the dummy coefficients on mispricing signal Q5 average to a positive number in nine (11) out of 12 calendar months for Specification 1 (Specifications 3 and 5) and are positive in 60% of the 310-month sample period. The performance of the strategy is also not statistically different between firms that announce or do not announce earnings in a given month. Fama-MacBeth coefficients on the product of an earnings announcement month dummy and the mispricing signal are negligible when the specifications add this variable and an earnings announcement month dummy to the regression. Coefficients on the mispricing signal remain of similar magnitude to those reported in Table 3.

Table 3
Fama-MacBeth cross-sectional regressions.

The table shows average coefficients and test statistics from Fama and MacBeth (1973) regressions of monthly returns on stock characteristics. Across different specifications, returns are regressed against end-of-prior-month values for the mispricing signal *M*, market beta, book-to-market, market capitalization, short-term reversal, momentum, long-term reversal, accruals, standardized unexpected earnings (SUE), gross profitability, and earnings yield. Columns under the ordinary least squares ("OLS") heading report results for signals from OLS regressions, and columns under the Theil (1950) and Sen (1968) ("TS") heading show results for signals from Theil–Sen regressions as described in the text. The table employs quintile dummies for the characteristics as regressors. Each month's quintiles are determined from sorts of firms with non-missing values for all characteristics. Size quintiles are based on NYSE breakpoints. The regressions include dummy variables for Quintiles 2, 3, 4, and 5 of each characteristic but the table displays only the coefficients of the quintile dummy with the largest amount of the characteristic (Q5) for brevity. Panel A shows results for the full-sample period, and Panel B shows results for the 1993–2012 subperiod. All regressions include industry dummy variables based on the 38 Fama and French industry classifications. The table also shows the average number of observations and average adjusted *R*-squared. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1987 to December 2012. All variables are defined in Appendix B.

	OLS										TS	
	Specification 1		Specification 2		Specification 3		Specification 4		Specification 5		Specification 6	
	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]
Panel A: Regressions with quintile dummies for full-sample period												
Mispricing Signal (M) (Q5)	0.4614 ***	[2.79]			0.5376 ***	[4.37]			0.3621 ***	[2.82]	0.4353 ***	[3.67]
Beta (Q5)			−0.1028	[−0.47]	−0.1274	[−0.60]	−0.0141	[−0.07]	−0.0593	[−0.30]	−0.0073	[−0.04]
Market capitalization (Q5)			−0.0248	[−0.12]	−0.0374	[−0.18]	−0.0536	[−0.27]	−0.1257	[−0.62]	−0.0173	[−0.08]
Book/market (Q5)			0.3022 *	[1.78]	0.1040	[0.62]	0.3552 **	[2.22]	0.2429	[1.45]	0.1818	[1.06]
Short-term reversal (Q5)			−1.1099 ***	[−6.24]	−1.0818 ***	[−6.14]	−1.1857 ***	[−6.84]	−1.1663 ***	[−6.79]	−1.1656 ***	[−6.74]
Momentum (Q5)			0.7910 ***	[3.75]	0.8079 ***	[3.81]	0.5447 ***	[2.67]	0.5627 ***	[2.76]	0.5746 ***	[2.82]
Long-term reversal (Q5)			−0.2791 **	[−2.46]	−0.3082 ***	[−2.75]	−0.2095 *	[−1.96]	−0.2063 *	[−1.94]	−0.2274 **	[−2.14]
Accruals (Q5)							−0.6624 ***	[−7.64]	−0.6498 ***	[−7.53]	−0.6400 ***	[−7.36]
SUE (Q5)							0.4094 ***	[4.31]	0.4043 ***	[4.25]	0.4138 ***	[4.35]
Gross profitability (Q5)							0.5516 ***	[5.00]	0.5457 ***	[4.94]	0.5265 ***	[4.76]
Earnings yield (Q5)							0.4754 ***	[4.32]	0.3732 ***	[3.22]	0.3220 ***	[2.79]
Intercept	0.2945	[0.77]	0.7835	[1.60]	0.6402	[1.27]	0.5373	[1.02]	0.5208	[0.97]	0.4131	[0.77]
Number of observations	1,349		1,349		1,349		1,349		1,349		1,349	
Adj. *R*-squared	0.041		0.073		0.074		0.078		0.080		0.079	
Industry control	Yes		Yes		Yes		Yes		Yes		Yes	
Panel B: Regressions with quintile dummies 1993–2012												
Mispricing Signal (M) (Q5)	0.5650 ***	[2.94]			0.5755 ***	[4.10]			0.4784 ***	[3.32]	0.4551 ***	[3.35]
Beta (Q5)			−0.1313	[−0.50]	−0.1657	[−0.65]	−0.0397	[−0.16]	−0.0896	[−0.38]	−0.0441	[−0.18]
Market capitalization (Q5)			−0.0906	[−0.37]	−0.1618	[−0.65]	−0.1252	[−0.55]	−0.2207	[−0.93]	−0.0942	[−0.39]
Book/market (Q5)			0.3941 *	[1.96]	0.1830	[0.92]	0.4597 **	[2.48]	0.3179	[1.62]	0.2754	[1.37]
Short-term reversal (Q5)			−1.0877 ***	[−5.01]	−1.0524 ***	[−4.89]	−1.1629 ***	[−5.50]	−1.1334***	[−5.42]	−1.1365 ***	[−5.38]
Momentum (Q5)			0.6287 **	[2.39]	0.6529 **	[2.47]	0.4480 *	[1.77]	0.4714 *	[1.86]	0.4826 *	[1.90]
Long-term reversal (Q5)			−0.3022 **	[−2.26]	−0.3281 **	[−2.50]	−0.2111 *	[−1.69]	−0.2075 *	[−1.67]	−0.2356 *	[−1.89]
Accruals (Q5)							−0.5721 ***	[−5.75]	−0.5589 ***	[−5.65]	−0.5485 ***	[−5.48]
SUE (Q5)							0.3112 ***	[2.73]	0.3072 ***	[2.69]	0.3130 ***	[2.75]
Gross profitability (Q5)							0.5637 ***	[4.26]	0.5513 ***	[4.18]	0.5380 ***	[4.06]
Earnings yield (Q5)							0.4074 ***	[3.16]	0.2818 **	[2.08]	0.2576 *	[1.94]
Intercept	0.1265	[0.29]	0.7091	[1.20]	0.5982	[0.98]	0.4982	[0.78]	0.4758	[0.73]	0.3683	[0.57]
Number of observations	1,437		1,437		1,437		1,437		1,437		1,437	
Adj. *R*-squared	0.045		0.079		0.080		0.084		0.086		0.085	
Industry control	Yes		Yes		Yes		Yes		Yes		Yes	

estly weaker momentum and earnings yield effects for the last 20 years.

4.3. 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 with 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.

Denote $r_{q,t+1}$ to be the industry-adjusted month $t+1$ return on quintile portfolio q based on $M_{j,t}$. With L factors, we estimate its alpha as the intercept in the time series regression

$$r_{q,t+1} = \alpha_q + \sum_{l=1}^{L} \beta_{q,l} F_{l,t+1} + \epsilon_{q,t+1} \tag{3}$$

where $F_{l,t+1}$ is the return difference (or excess return) of the lth 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 Quintile 1 portfolios, a metric of the mispricing signal's ability to earn abnormal profits, should be significantly positive.

The Table 4 industry-adjusted returns are essentially a zero-factor specification. The Panel A industry-adjusted 48 basis points per month spread between mispricing Quintiles 5 and 1 is not identical to the 46 basis point spread in Table 3, Panel A. The spreads differ because Table 4 lifts the requirement that firms possess data for all of the Table 3 specifications. Table 4 also adjusts for industry effects by subtracting the industry return from the dependent variable, while Table 3 employs industry dummies as regressors. The industry-adjusted returns are monotonic. Annualized Sharpe ratios here, and throughout the paper, are obtained by multiplying the t-statistics of the intercepts by 0.20, the square root of the ratio of 12 to the number of time series observations (typically, 310). For Panel A, the Sharpe ratios of the Quintile 5–1 spreads range from 0.63 (zero-factor model) to 1.01 (six-factor model) for OLS-based signals, and TS signals have Sharpe ratios of up to 1.63. Our tables omit these ratios for brevity.

The Table 4 six- and eight-factor specifications nest the widely used Fama and French (1993) three-factor, Carhart (1997) four-factor, and Fama and French (2014) five-factor models within them. The six-factor model contains the market excess return (Mkt_RF), a size factor (SMB, small minus big), a value factor (HML, high minus low), momentum (Mom), a short-term reversal factor (ST_Rev), and a long-term reversal factor (LT_Rev). The Table 4 eight-factor model also employs an investment factor (CMA, conservative minus aggressive) and profitability factor (RMW, robust minus weak) and represents the broadest factor model available in the Kenneth French data library.[22] Appendix B provides more detail on all of the factors used in our analysis.

The six factor betas in the Table 4 six-factor model, all similar to their eight-factor counterparts, indicate that our mispricing strategy is exposed to five of the model's six dimensions of factor risk. Compared with overpriced firms, underpriced firms are more exposed to the returns of value firms with low betas and poor short- and medium-term returns. These marginal effects are broadly similar to those in Table 1 when studying characteristics across the mispricing quintiles. Underpriced firms also are significantly more exposed to the returns of firms with better long-term returns, largely because of HML's correlation with the long-term past returns factor, noted by Gerakos and Linnainmaa (2017).

The Table 4 alphas take out the return contribution of these factor exposures. The fairly monotonic alphas of the six- and eight-factor models in Panel A are similar. The monotonicity strengthens the argument that fundamental analysis works. About 60 basis points per month distinguish the two extreme quintile portfolios' OLS alphas, with almost all of the profitability of the alpha spreads coming from the most overpriced quintiles, which is consistent with short-selling constraints providing more investment opportunities on the short side.[23] These spreads are also larger than the HML premium, even after controlling for HML. For signals from Theil-Sen regressions, all spreads are economically and statistically even more significant, ranging from 74 (zero-factor model) to 87 basis points (six-factor model) per month.

Panel B weights returns by market capitalization as of the end of month t, which is prior to the return month. At 47 and 39 basis points per month, these value-weighted OLS six- and eight-factor alpha spreads are weaker than their equally weighted counterparts but are still significant. The same is true for the Panel B TS estimates. Depending on the estimation technique and factor adjustment, value weighting tends to reduce alphas by one quarter to one third of their equal-weighted counterparts. 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, lower spreads for value weighting could also be an artifact of poor diversification. Value-weighted portfolios containing large firms present special inference problems when they contain a few large firms.[24] The existence of these firms makes portfolio alpha estimates imprecise and largely determined by the firm-specific return realizations of a few large firms rather than the portfolio's true alpha. For example, with less than 1% of its stocks from

[22] 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).

[23] Consequently, the risk-adjusted quintile spreads could be weaker when excluding stocks that were on special and, hence, difficult to borrow and sell short or when accounting for stock lending fees (Beneish, Lee, and Nichols, 2015).

[24] This issue is not present in a value-weighted index of all stocks, which achieves diversification by having many large firms.

Table 4
Factor model time series regressions.

The table shows average industry-adjusted portfolio returns (as described in the text), as well as intercepts, slope coefficients, and *t*-statistics from time series regressions of monthly industry-adjusted portfolio returns on six or eight factors. 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. The table reports averages and regression statistics separately for each of the five portfolios, Q1–Q5, and for the corresponding times series of return spreads between the most undervalued (Q5) and overvalued (Q1) stock quintiles. For the six-factor model, the factors are Mkt_RF (market excess return), SMB (small minus big), HML (high minus low), Mom (momentum), ST_Rev (short-term reversal factor) and LT_Rev (long-term reversal factor), obtained from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The eight-factor model also includes the CMA (conservative minus aggressive) and RMW (robust minus weak) factors from the Kenneth French data library. Columns under the ordinary least squares ("OLS") heading report results for signals from OLS regressions, and columns under the Theil (1950) and Sen (1968) ("TS") heading show results for signals from Theil-Sen regressions as described in the text. The table also shows the number of observations and adjusted *R*-squared. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1987 to December 2012. All variables are defined in Appendix B.

	OLS												TS	
	Q1 (overvalued)		Q2		Q3		Q4		Q5 (undervalued)		Q5–Q1 (undervalued–overvalued)		Q5–Q1 (undervalued–overvalued)	
	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]
Panel A: Equal-weighted portfolios (with industry control)														
Industry-adjusted returns	−0.4482***	[−3.36]	−0.1884	[−1.30]	−0.0852	[−0.62]	0.0091	[0.08]	0.0332	[0.30]	0.4814***	[3.19]	0.7400***	[4.97]
Six-factor model														
Alpha	−0.5978***	[−5.34]	−0.2976**	[−2.48]	−0.1980	[−1.62]	−0.0764	[−0.68]	0.0253	[0.26]	0.6232***	[5.11]	0.8709***	[8.31]
Mkt_RF	0.0805***	[2.95]	0.0838***	[2.86]	0.0792***	[2.66]	0.0635**	[2.32]	−0.0290	[−1.22]	−0.1095***	[−3.68]	−0.1784***	[−6.98]
SMB	−0.0708*	[−1.72]	−0.1660***	[−3.75]	−0.1632***	[−3.62]	−0.0873**	[−2.11]	−0.1126***	[−3.13]	−0.0418	[−0.93]	−0.1183***	[−3.07]
HML	−0.0042	[−0.10]	−0.0026	[−0.06]	0.1368***	[2.99]	0.2215***	[5.28]	0.2644***	[7.24]	0.2685***	[5.89]	0.3241***	[8.27]
Mom	0.2681***	[11.17]	0.2688***	[10.43]	0.2069***	[7.88]	0.1004***	[4.17]	0.0393*	[1.87]	−0.2288***	[−8.75]	−0.2136***	[−9.50]
ST_Rev	−0.0895***	[−2.78]	−0.1277***	[−3.70]	−0.0432	[−1.23]	−0.0164	[−0.51]	−0.0256	[−0.91]	0.0639*	[1.82]	0.0834***	[2.77]
LT_Rev	−0.0328	[−0.62]	−0.0985*	[−1.74]	−0.1670***	[−2.89]	−0.1654***	[−3.13]	−0.1679***	[−3.64]	−0.1350**	[−2.35]	−0.0357	[−0.72]
R-squared	0.35		0.37		0.26		0.17		0.30		0.40		0.54	
Number of observations	310		310		310		310		310		310		310	
Eight-factor model														
Alpha	−0.6840***	[−6.04]	−0.4002***	[−3.31]	−0.3469***	[−2.95]	−0.2246**	[−2.08]	−0.0707	[−0.71]	0.6133***	[4.83]	0.8039***	[7.43]
Mkt_RF	0.0998***	[3.49]	0.1072***	[3.51]	0.1120***	[3.77]	0.0987***	[3.62]	−0.0028	[−0.11]	−0.1026***	[−3.20]	−0.1593***	[−5.84]
SMB	−0.0204	[−0.48]	−0.1069**	[−2.38]	−0.0746*	[−1.71]	−0.0049	[−0.12]	−0.0667*	[−1.81]	−0.0464	[−0.98]	−0.0880**	[−2.19]
HML	−0.0716	[−1.31]	−0.0843	[−1.45]	0.0223	[0.39]	0.0989*	[1.90]	0.1739***	[3.65]	0.2455***	[4.02]	0.2585***	[4.97]
Mom	0.2517***	[10.53]	0.2494***	[9.77]	0.1784***	[7.19]	0.0727***	[3.19]	0.0223	[1.07]	−0.2294***	[−8.56]	−0.2252***	[−9.87]
ST_Rev	−0.1022***	[−3.23]	−0.1422***	[−4.21]	−0.0662**	[−2.01]	−0.0351	[−1.16]	−0.0324	[−1.17]	0.0699**	[1.97]	0.0799***	[2.64]
LT_Rev	0.0260	[0.47]	−0.0315	[−0.53]	−0.0604	[−1.05]	−0.0790	[−1.49]	−0.1376***	[−2.83]	−0.1636***	[−2.63]	−0.0202	[−0.38]
CMA	−0.0172	[−0.22]	−0.0124	[−0.15]	−0.0423	[−0.52]	0.0090	[0.12]	0.0720	[1.04]	0.0891	[1.01]	0.0648	[0.86]
RMW	0.2291***	[4.18]	0.2681***	[4.58]	0.4037***	[7.10]	0.3712***	[7.10]	0.2009***	[4.19]	−0.0282	[−0.46]	0.1314**	[2.51]
R-squared	0.39		0.41		0.37		0.29		0.34		0.40		0.55	
Number of observations	310		310		310		310		310		310		310	

(*continued on next page*)

Table 4 (*continued*)

	OLS												TS	
	Q1 (overvalued)		Q2		Q3		Q4		Q5 (undervalued)		Q5–Q1 (undervalued–overvalued)		Q5–Q1 (undervalued–overvalued)	
	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]
Panel B: Value-weighted portfolios (with industry control)														
Industry-adjusted returns	−0.4878**	[−2.48]	−0.2453	[−1.07]	−0.1999	[−0.95]	−0.0025	[−0.02]	0.0277	[0.19]	0.5155***	[2.73]	0.5966**	[2.40]
Six-factor model														
Alpha	−0.4594***	[−3.09]	−0.2573*	[−1.79]	−0.2125	[−1.50]	−0.0448	[−0.34]	0.0066	[0.05]	0.4660***	[2.87]	0.6417***	[3.96]
Mkt_RF	0.1031***	[2.85]	0.0351	[1.00]	0.0389	[1.12]	0.0983***	[3.03]	0.1000***	[3.21]	−0.0031	[−0.08]	−0.2128***	[−5.39]
SMB	−0.6386***	[−11.68]	−0.8843***	[−16.72]	−0.7925***	[−15.14]	−0.4358***	[−8.89]	−0.3154***	[−6.70]	0.3232***	[5.40]	0.1029*	[1.72]
HML	−0.1230**	[−2.22]	0.1233**	[2.30]	0.1462***	[2.75]	0.1649***	[3.31]	0.2007***	[4.20]	0.3237***	[5.33]	0.6810***	[11.25]
Mom	0.2016***	[6.33]	0.2450***	[7.95]	0.1873***	[6.15]	0.0911***	[3.19]	−0.0051	[−0.19]	−0.2067***	[−5.93]	−0.3798***	[−10.93]
ST_Rev	−0.1567***	[−3.67]	−0.1785***	[−4.32]	−0.0261	[−0.64]	−0.0361	[−0.94]	−0.0193	[−0.52]	0.1374***	[2.94]	0.3034***	[6.51]
LT_Rev	−0.0968	[−1.38]	−0.0072	[−0.11]	−0.1087	[−1.62]	−0.1092*	[−1.74]	−0.1052*	[−1.75]	−0.0084	[−0.11]	0.0006	[0.01]
R-squared	0.47		0.64		0.58		0.34		0.26		0.31		0.61	
Number of observations	310		310		310		310		310		310		310	
Eight-factor model														
Alpha	−0.5241***	[−3.40]	−0.4394***	[−3.19]	−0.3746***	[−2.81]	−0.2133*	[−1.67]	−0.1334	[−1.06]	0.3907**	[2.34]	0.3070**	[2.04]
Mkt_RF	0.1208***	[3.11]	0.0763**	[2.19]	0.0688**	[2.05]	0.1368***	[4.23]	0.1322***	[4.16]	0.0114	[0.27]	−0.1150***	[−3.02]
SMB	−0.6079***	[−10.63]	−0.7787***	[−15.23]	−0.6835***	[−13.84]	−0.3387***	[−7.13]	−0.2352***	[−5.04]	0.3727***	[6.01]	0.2485***	[4.44]
HML	−0.1843**	[−2.49]	−0.0208	[−0.31]	0.0403	[0.63]	0.0307	[0.50]	0.0885	[1.46]	0.2728***	[3.39]	0.3442***	[4.75]
Mom	0.1902***	[5.85]	0.2105***	[7.24]	0.1546***	[5.51]	0.0592**	[2.19]	−0.0315	[−1.19]	−0.2217***	[−6.29]	−0.4369***	[−13.74]
ST_Rev	−0.1611***	[−3.75]	−0.2046***	[−5.32]	−0.0602	[−1.62]	−0.0599*	[−1.67]	−0.0387	[−1.10]	0.1224***	[2.62]	0.2902***	[6.89]
LT_Rev	−0.0770	[−1.02]	0.1138*	[1.68]	0.0506	[0.78]	0.0009	[0.01]	−0.0154	[−0.25]	0.0616	[0.75]	0.0571	[0.77]
CMA	0.0501	[0.47]	−0.0275	[−0.29]	−0.1576*	[−1.70]	−0.0206	[−0.23]	−0.0128	[−0.15]	−0.0629	[−0.54]	0.3775***	[3.59]
RMW	0.1344*	[1.80]	0.4791***	[7.19]	0.5060***	[7.86]	0.4405***	[7.11]	0.3634***	[5.97]	0.2290***	[2.83]	0.6248***	[8.57]
R-squared	0.48		0.69		0.66		0.44		0.35		0.33		0.69	
Number of observations	310		310		310		310		310		310		310	

the top NYSE size decile, accurately estimating the value-weighted mean return or alpha of the most underpriced quintile is impossible. Typically, only four of these mega-cap firms, often one hundred times larger than the typical firm, appear in the most underpriced quintile.[25]

In all cases, the alphas of the most underpriced quintile of stocks insignificantly differ from zero. Hence, like most efficient market anomalies in finance, the factor-adjusted profitability of spread portfolio strategies mostly come from the short leg of the strategy. Short-sales constraints can hinder the ability to implement the spread strategy and short-sales costs can reduce its profitability. However, other explanations exist for the absence of significantly positively alphas.

5. Convergence to fair value better explains the results than alternatives

Underpriced stocks have lower market betas and higher average returns than overpriced stocks. Moreover, other potential risk factors are unlikely to account for our trading strategy's profitability.

5.1. Alpha spreads already control for known sources of risk and firm size

We regress the time series of signal Q5 regression coefficients from Table 3, Panel A (which Fama and MacBeth note are portfolio returns), on the factors in Table 4. The intercept (alpha) negligibly differs from the coefficients in Table 3, Panel A, and factor betas are mostly negligible. Thus, the Table 3 controls for other characteristics largely eliminate the factor risks that the Table 4 alphas already take into account.[26]

We also ran regressions of each mispricing quintile's industry-adjusted return against the Stambaugh and Yu (2017) mispricing factor used in their three-factor model. The betas against the mispricing factor of the extreme mispricing quintiles were mostly smaller at 0.181 (Q1) and 0.205 (Q5) than the betas of the three interior quintiles, which were 0.268 (Q2), 0.263 (Q3), and 0.194 (Q4), respectively. Moreover, the Q5–Q1 industry-adjusted return spread has a mispricing factor beta of only 0.024. By contrast, the spread portfolios represented by HML and Mom have mispricing factor betas of 0.277 and 0.671, respectively. Thus, an omitted mispricing factor does not account for any of our results.

Berk (1995) notes that firms with lower market capitalizations (like our underpriced firms) tend to have higher discount rates for future dividend streams, other things equal. Irrespective of whether small firms' higher discount rates are driven by omitted risk factors or animal spirits, we can assess the degree to which size differences across the mispricing quintiles play a role in our findings. If our mispricing signal works because it proxies for an omitted risk factor linked to market capitalization, six- and eight-factor models (which contain size factors) should generate lower abnormal return spreads than spreads without factor controls. However, the six- and eight-factor alpha spreads in industry-adjusted returns, which control for known sources of factor risk, including size, exceed the (top row) spread in industry-adjusted returns.

Finally, if an omitted risk variable tied to cross-sectional differences in size explains our alpha spreads, stale mispricing signals should produce almost the same alpha spreads as fresh signals. 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 prior 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 month $t-1$ is 0.90, and the same correlation for the book-to-market ratio is 0.97. Moreover, the rank correlation between months t and $t-12$ is 0.55 for our mispricing measure, and it is 0.79 for the book-to-market ratio. If omitted risk accounts for our alpha spreads, then that risk exposure would have to change rapidly, instead of being due to a more stable characteristic, such as the cross-sectional difference in firm size.

5.2. Signal delay

Using (for fair comparisons) returns beginning in September 1990, Fig. 1 graphs the six-factor (Panel A) and eight-factor (Panel B) 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 zero to 35 months.[27] The decay in the signal's efficacy is rapid in the first month. For example, with the six-factor alpha, the signal's initial ability to earn abnormal returns of 73 basis points over the next month drops to 46 basis points when the mispricing signal is one month old. While strategy performance tends to decrease with more delay, the pattern is not monotonic, as signals lagged by three and nine months still generate alphas of 57 basis points per month. For signals between two and 12 months old, the signal generates six-factor alphas of 47 basis points per month on average. Alphas average 20 basis points per month for signals that are one to two years old. The eight-factor alpha pattern is highly similar.

Fig. 1 also shows the strategy's performance when constructing a mispricing signal using accounting data that are lagged by six months relative to the point date

[25] 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 alpha of the 20% most underpriced stocks overall exceeds that of the 20% most overpriced by 101 basis points per month when benchmarked against the eight-factor model. Yet, spreads of this magnitude, despite being rare, fail to attain the 5% significance threshold. The insignificance stands in contrast to the significant but smaller negative alpha of the far larger number of overpriced firms in the same size quintile.

[26] Aretz, Bartram, and Pope (2010) show that book-to-market, size, and momentum capture cross-sectional variation in exposures to a broad set of macroeconomic factors identified in the prior literature as potentially important for pricing equities.

[27] Because of the later start date, the alphas for the zero lag differ slightly from those in Table 4, Panel A.

Fig. 1. Signal decay. The figure shows portfolio alphas from 36 pairs of six- and eight-factor model time series regressions. Signals are constructed using alternatively point-in-time dates (solid line) or point-in-time dates plus six months (dashed line) for the timing of the availability of accounting data. Each month, stocks are sorted into quintiles (Q1–Q5) based on a lagged mispricing signal (*M*), for lags from zero to 35 months, and combined into equally weighted portfolios. Each spread portfolio return (in excess of the industry portfolios based on the 38 Fama and French industry classifications) from one of the 36 signals, the difference between the returns of portfolios Q5 and Q1, is regressed on a set of factors. For the six-factor model, the factors are Mkt_RF (market excess return), SMB (small minus big), HML (high minus low), Mom (momentum), ST_Rev (short-term reversal factor) and LT_Rev (long-term reversal factor), obtained from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The eight-factor model also includes the CMA (conservative minus aggressive) and RMW (robust minus weak) factors from the Kenneth French data library. The sample consists of all ordinary common stocks of US 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 $5. The sample period is September 1990 to December 2012. All variables are defined in Appendix B.

(dashed lines). Because the point date is after the fiscal close date, lagging point-in-time data is more conservative than employing similar lags using the regular Compustat database. With our more conservative six-month delay, alpha spreads shrink by 20–25 basis points, but are still significant.

Fig. 2 shows the six- and eight-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 a quarter. While both the stale and the most recent coefficients are estimated with error, averaging the weights over various windows does not enhance performance or prevent performance decay with a delayed signal.

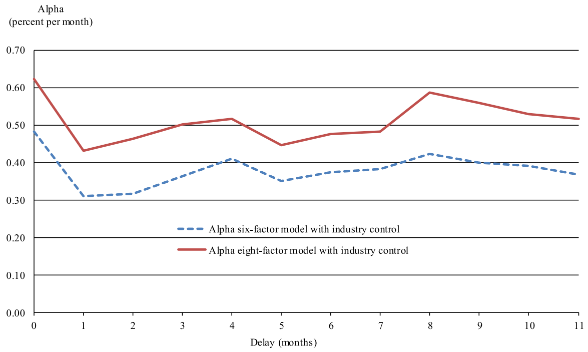

Fig. 2. Accounting weights. The figure shows alphas from 12 pairs of factor model time series regressions. Stocks are sorted each month into quintiles (Q1–Q5) based on a mispricing signal (*M*) and combined into equally weighted portfolios. The mispricing signal is based on fair value estimates, derived from cross-sectional regressions that weight accounting variables. The fair value prediction that determines the five quintile portfolios uses coefficients that are from fair value regressions lagged between zero and 11 months along with the accounting variables from lag zero. Each spread portfolio return (in excess of the industry portfolios based on the 38 Fama and French industry classifications) from one of the 12 signals, the difference between the returns of portfolios Q5 and Q1, is regressed on a set of factors. For the six-factor model, the factors are Mkt_RF (market excess return), SMB (small minus big), HML (high minus low), Mom (momentum), ST_Rev (short-term reversal factor) and LT_Rev (long-term reversal factor), obtained from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The eight-factor model also includes the CMA (conservative minus aggressive) and RMW (robust minus weak) factors from the Kenneth French data library. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1988 to December 2012. All variables are defined in Appendix B.

The signal delay results also help to 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 zero to 11 months. The average alphas from such a relaxed strategy, as measured by averaging the first 12 alphas in Fig. 1, namely, 49 and 44 basis points per month for the six- and eight-factor models, respectively, stem from trades with 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 51% per month, whereas holding positions for one year leads to monthly turnover of 8%, with both turnover ratios almost equally split between the long and short positions. The relaxed strategy's turnover thus requires unrealistically high trading costs before such costs offset the alpha spread (for an estimate of U.S. transactions costs see Bartram and Grinblatt (2017). We also verify that the relaxed strategy's alpha from a one-year holding period is significant using the Jegadeesh and Titman (1993, 2001) technique for estimation of the test statistics.

Table 5
Mispricing strategies within quintiles of other anomalies.
The table shows intercepts and *t*-statistics from time series regressions of monthly industry-adjusted portfolio returns of a mispricing-based spread portfolio on six (Panel A) or eight (Panel B) factors. Stocks are first sorted each month into quintiles, designated by column heading, based on the row's firm characteristic. Within each of the former quintiles, stocks are further sorted into quintiles based on the mispricing signal and combined into equally weighted portfolios. The industry-adjusted return difference of the most underpriced and overpriced stocks within each cell are then regressed on six and eight factors. For the six-factor model, the factors are Mkt_RF (market excess return), SMB (small minus big), HML (high minus low), Mom (momentum), ST_Rev (short-term reversal factor) and LT_Rev (long-term reversal factor), obtained from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The eight-factor model also includes the CMA (conservative minus aggressive) and RMW (robust minus weak) factors from the Kenneth French data library. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1987 to December 2012. All variables are defined in Appendix B.

	Q1		Q2		Q3		Q4		Q5	
Variable	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]
Panel A: Six-factor alphas of long-short mispricing strategies										
Beta	0.7066***	[4.56]	0.3736**	[2.58]	0.5126***	[3.56]	0.4138**	[2.40]	0.7525***	[4.04]
Book/market	0.4123*	[1.95]	0.2349	[1.46‘]	0.4028**	[2.57]	0.4948***	[3.11]	0.4846***	[2.89]
Market capitalization	0.8621***	[5.30]	0.7783***	[4.80]	0.7415***	[4.80]	0.5527***	[3.80]	0.2656*	[1.96]
Short-term reversal	1.1474***	[5.65]	0.7717***	[5.07]	0.5141***	[3.21]	0.0916	[0.61]	0.1723	[0.92]
Momentum	1.1044***	[5.82]	0.6540***	[3.99]	0.4950***	[3.37]	0.5871***	[3.82]	0.7501***	[3.83]
Long-term reversal	0.6473***	[3.38]	0.7433***	[4.62]	0.5243***	[3.44]	0.4824***	[3.03]	0.8142***	[4.04]
Accruals	0.6504***	[3.64]	0.6358***	[4.26]	0.6084***	[3.95]	0.5585***	[3.19]	0.5152**	[2.24]
SUE	0.6647***	[3.52]	0.5426***	[3.48]	0.4615***	[3.06]	0.9610***	[5.94]	0.8052***	[4.88]
Gross profitability	0.5798***	[3.31]	0.6178***	[3.78]	0.5031***	[3.17]	0.6362***	[3.80]	0.4306**	[2.33]
Scaled NOA	1.1401***	[5.34]	0.4899***	[2.97]	0.3133*	[1.94]	0.5988***	[3.71]	0.6507***	[3.75]
Share issuance	0.4006***	[2.81]	0.5818***	[4.08]	0.4666***	[2.96]	0.6528***	[3.66]	0.5649***	[2.93]
Asset growth	0.5854***	[3.40]	0.3891***	[2.70]	0.7570***	[5.47]	0.7354***	[4.23]	0.4320**	[2.17]
Capital investment	0.5611***	[3.17]	0.3701**	[2.20]	0.3779**	[2.38]	0.8187***	[5.24]	0.6171***	[3.39]
Investment ratio	0.5025***	[3.06]	0.5903***	[3.59]	0.7694***	[5.15]	0.7517***	[4.50]	0.6091***	[2.99]
External financing	0.4274***	[2.78]	0.3787**	[2.31]	0.5534***	[3.38]	0.5600***	[3.11]	0.4572**	[2.07]
Z-score	0.4821**	[2.47]	0.3137**	[2.07]	0.4994***	[2.93]	0.5599***	[3.24]	0.6645***	[3.26]
Leverage	0.6900***	[3.24]	0.6642***	[3.96]	0.2939*	[1.93]	0.3526**	[2.43]	0.6085***	[3.73]
Illiquidity	0.6265***	[4.67]	0.7351***	[5.07]	0.7322***	[4.72]	0.5024***	[2.92]	0.7572***	[3.88]
Earnings yield	0.4072**	[2.07]	0.4904***	[2.70]	0.4286***	[2.79]	0.3555***	[2.60]	0.4838***	[3.07]
Dividend/price	0.7381***	[4.77]	0.5341***	[3.01]	0.6333***	[3.89]	0.4584***	[2.98]	0.5531***	[3.09]
Cash flow/price	0.6512***	[3.39]	0.2443	[1.39]	0.1678	[1.20]	0.5626***	[3.64]	0.5413***	[3.32]
V/P	0.6256***	[3.17]	0.5141***	[3.07]	0.4796***	[3.31]	0.3204**	[2.37]	0.5765***	[3.61]
Panel B: Eight-factor alphas of long-short mispricing strategies										
Beta	0.6886***	[4.27]	0.3097**	[2.07]	0.4516***	[3.05]	0.2862	[1.64]	0.6512***	[3.40]
Book/market	0.3378	[1.55]	0.2685	[1.60]	0.5072***	[3.16]	0.5523***	[3.43]	0.6487***	[3.87]
Market capitalization	0.8367***	[4.98]	0.6307***	[3.87]	0.5310***	[3.51]	0.4308***	[2.89]	0.0996	[0.73]
Short-term reversal	1.1781***	[5.57]	0.7496***	[4.76]	0.4981***	[3.00]	0.1228	[0.79]	0.1553	[0.80]
Momentum	1.1272***	[5.73]	0.7154***	[4.24]	0.5269***	[3.47]	0.5509***	[3.44]	0.8120***	[3.99]
Long-term reversal	0.6404***	[3.21]	0.7238***	[4.37]	0.5980***	[3.79]	0.4732***	[2.85]	0.7741***	[3.69]
Accruals	0.6838***	[3.68]	0.6609***	[4.29]	0.5600***	[3.55]	0.5075***	[2.79]	0.5488**	[2.29]
SUE	0.7341***	[3.74]	0.4388***	[2.74]	0.4129***	[2.63]	0.9553***	[5.69]	0.8121***	[4.73]
Gross profitability	0.6050***	[3.31]	0.6803***	[4.03]	0.5355***	[3.26]	0.5844***	[3.37]	0.4404**	[2.29]
Scaled NOA	0.9964***	[4.52]	0.4326**	[2.53]	0.2368	[1.41]	0.6645***	[3.99]	0.6660***	[3.73]
Share issuance	0.5324***	[3.70]	0.6795***	[4.68]	0.4184**	[2.56]	0.5655***	[3.05]	0.5921***	[2.95]
Asset growth	0.5712***	[3.18]	0.3960***	[2.66]	0.7381***	[5.16]	0.7528***	[4.16]	0.4766**	[2.30]
Capital investment	0.5835***	[3.16]	0.4195**	[2.41]	0.4096**	[2.49]	0.8340***	[5.18]	0.6116***	[3.22]
Investment ratio	0.4909***	[2.90]	0.5186***	[3.03]	0.7951***	[5.11]	0.7865***	[4.53]	0.6880***	[3.25]
External financing	0.4602***	[2.91]	0.4708***	[2.80]	0.5945***	[3.49]	0.5369***	[2.87]	0.4686**	[2.04]
Z-score	0.5641***	[2.78]	0.4070***	[2.62]	0.6390***	[3.71]	0.6327***	[3.61]	0.7241***	[3.50]
Leverage	0.6510***	[2.95]	0.6288***	[3.60]	0.3592**	[2.31]	0.3126**	[2.10]	0.7214***	[4.34]
Illiquidity	0.5869***	[4.21]	0.5666***	[3.86]	0.6655***	[4.17]	0.3623**	[2.14]	0.7479***	[3.68]
Earnings yield	0.5953***	[2.97]	0.5178***	[2.75]	0.5627***	[3.69]	0.4145***	[3.04]	0.5812***	[3.72]
Dividend/price	0.7528***	[4.68]	0.5842***	[3.18]	0.6962***	[4.16]	0.5296***	[3.36]	0.6060***	[3.30]
Cash flow/price	0.8311***	[4.25]	0.2803	[1.54]	0.2665*	[1.87]	0.6796***	[4.46]	0.6124***	[3.72]
V/P	0.6177***	[3.00]	0.5105***	[2.93]	0.5016***	[3.37]	0.3357**	[2.46]	0.6082***	[3.81]

6. What's in the black box?

To understand whether the signal's black box proxies for an anomaly previously seen in the literature, Table 5 investigates how the mispricing signal relates to a set of 22 known anomalies. It reports the six- (Panel A) and eight-factor (Panel B) alphas of trading strategies formed from the mispricing signal. The trades of these strategies take place within 110 subgroups of stocks that share similar amounts of an alternative characteristic known to predict returns and alphas. These include the characteristics already considered in Table 3, i.e., beta, book-to-market, market capitalization, short-term reversal, momentum, long-term reversal, accruals, earnings surprise, gross profitability, and earnings yield. Further predictors are scaled net operating assets (NOA), share issuance, asset growth, capital investment, investment ratio, external financing, Z-score, leverage, illiquidity, dividend/price, cash flow/price, and value/price (V/P) (Frankel and Lee, 1998). Sample size limitations make it undesirable to control for all characteristics simultaneously.

Each month, stocks are sorted into quintiles based on one of the 22 predictive characteristics. Within each quintile, they are then sorted into quintiles based only on our mispricing variable. The Table 5 alpha spreads are from a long-short trading strategy in the extreme mispricing quintiles produced by this sequential sort. If the 22 characteristics are highly related to our mispricing variable, the lack of mispricing signal variation in the sequential sort's second step should eliminate significant alpha spreads from the mispricing signal.

Apparent from both panels, statistically and economically significant alpha spreads exist in almost all the quintiles of the other characteristics. Of the 110 six-factor alphas in Panel A, five are scattered exceptions to significance and only one of them (short-term reversal) exists in an extreme quintile of the characteristic. Of the 110 eight-factor alphas in Panel B, eight are scattered exceptions to significance and only three of them (book-to-market, market capitalization, short-term reversal) are in an extreme quintile of the characteristic. Moreover, both panels' alphas control for factors tied to size, book-to-market, and past returns, and the eight-factor alphas also control for profitability and investment. Thus, correlations between the 22 characteristics and the mispricing signal are unlikely drivers of the alpha from trading on the mispricing signal.

Table 5 also indicates that the profitability of the mispricing strategy could be enhanced by focusing on groups of stocks with particular characteristics. These include stocks with small market capitalizations, low past returns (over short- and medium-term horizons), high earnings surprises, low scaled NOA, and low dividend-to-price. Mispricing estimates employing the staler market capitalizations at fiscal close (as opposed to those at the time of the mispricing signal) yield significant abnormal returns that are similar to those presented in the paper. This further rules out the issuance anomaly as a potential driver of our findings see, e.g., Ikenberry et al. (1995), Loughran and Ritter (1995), Mitchell and Stafford (2000), Teoh and Wong (2002), Schultz (2003), Daniel and Titman (2006), Fama and French (2008), and Pontiff and Woodgate (2008)).

Direct analysis of the alpha-generating role played by each of the 28 accounting items in the black box is another way to expose drivers of the mispricing signal's profitability. Unfortunately, no straightforward way exists to analyze the separate roles of each accounting item because they are highly collinear. A number of them are perfectly redundant in some or all months.[28] The marginal contribution of each variable to the strategy profitability of the 27 remaining accounting items is necessarily trivial. Table 6 tries to circumvent this thorny issue. It shows industry- and risk-adjusted performance (using the six- and eight-factor models) of long-short trading strategies for modified mispricing signals derived from alternative specifications of the peer-implied fair value regression. Its various specifications address how much each of the 28 accounting items contributed to overall performance by adding (Panel A) or subtracting (Panel B) the accounting items in a particular sequence, determined by coverage. Panel C looks at performance using only the 16 balance sheet items or only the 12 income statement and cash flow statement items to determine fair value.

Each of the 29 fair value regression specifications in Panel A uses the accounting item listed in its row plus all of the accounting items in the rows above as regressors. Each of the 29 specifications in Panel B use all accounting items excluding the accounting items in the rows above.[29] Thus, the Panel A starting point is the signal from monthly cross-sectional fair value regressions without any accounting variables. As we subsequently add each of the 28 items one by one, performance from the resulting signal tends to increase, though not entirely monotonically and sometimes in smaller and sometimes in larger increments. Performance turns statistically significant with the addition of the balance sheet item AOQH (other assets) and noticeably increases with the inclusion of SALEQH (sales/turnover) and IBQH (income before extraordinary items). The inclusion of IBQH nearly doubles performance. Adding the only cash flow variable DVQH (cash dividends) to the signal detracts from performance. The bottom of Panel A, as well as the top of Panel B, considers all 28 items from the balance sheet, income statement, and cash flow statement and was reported in Table 4. This analysis shows that our approach requires more than one accounting item to generate significant trading profits.

In Panel B, performance declines as items are dropped. The two largest declines in the six- and eight-factor alphas occur when PIQH (pretax income) is dropped from the specification (which still includes all of the accounting items below it as regressors) and when SALEQH (sales/turnover) is dropped from the fair value regression. The ten remaining items below, which have the least coverage among the 28, cannot produce significant trading profits without assistance from some of the items listed above them.

Signals from fair value regressions without any accounting variables (the top row of Panel A or the bottom row of Panel B) are effectively pseudo signals that capture relative market capitalization. Controlling for SMB and other standard risk factors, we no longer find risk-adjusted returns from this pseudo signal. Controlling for these risk factors, including SMB, has little effect on the return spreads of the signal with all 28 accounting variables. Consequently, the

[28] For example, included in the 28 accounting items are CEQQH (common/ordinary equity), PSTKQH (preferred/preference stock), and SEQQH (stockholders equity) with the last being the sum of the first two. Due to the collinearity of the regressions, the estimated coefficients can flip signs from month to month. To illustrate, the average coefficient of net income is 7.2, with a minimum of −219.9 and a maximum of 114.5. Nevertheless, this is not a concern because only the predicted value is necessary to construct the mispricing signal.

[29] For fair comparisons, we require firms to have non-missing data on all 28 accounting items even if we do not use all 28 items in all but one of the alternative fair value regressions.

Table 6
Signal additions and deletions.
The table shows average industry-adjusted portfolio returns, as well as intercepts and t-statistics from time series regressions of monthly industry-adjusted portfolio returns on six or eight factors. Each row uses alternative constructions of the mispricing signal that vary with the set of accounting items used to obtain fair value. In Panel A, the accounting items listed are sequentially added as regressors in the fair value regression. In Panel B, the accounting items listed are sequentially dropped from the fair value regression. Panel C shows results separately for fair value regressions with only the balance sheet items, and only the income and cash flow statement items, respectively. Stocks are sorted each month into quintiles based on the mispricing signal (M) and combined into equally weighted portfolios. The table reports averages and regression statistics for the corresponding times series of return spreads between the most undervalued (Q5) and overvalued (Q1) stock quintiles. For the six-factor model, the factors are Mkt_RF (market excess return), SMB (small minus big), HML (high minus low), Mom (momentum), ST_Rev (short-term reversal factor) and LT_Rev (long-term reversal factor), obtained from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The eight-factor model also includes the CMA (conservative minus aggressive) and RMW (robust minus weak) factors from the Kenneth French data library. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The sample consists of all ordinary common stocks of US 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 $5. The sample period is March 1987 to December 2012. All variables are defined in Appendix B.

	Industry-adjusted return		Six-factor alpha		Eight-factor alpha	
Variables	Coefficient	[t-statistic]	Coefficient	[t-statistic]	Coefficient	[t-statistic]
Panel A: Variable additions (sequentially added variables)						
None (just regression intercept)	−0.1172	[−0.70]	−0.1559	[−1.20]	−0.0718	[−0.54]
ATQH (total assets)	−0.0353	[−0.20]	−0.0946	[−0.70]	−0.0602	[−0.43]
SEQQH (total stockholders equity)	0.1970	[1.01]	0.1862	[1.28]	0.0993	[0.66]
ICAPTQH (total invested capital)	0.1884	[1.14]	0.1754	[1.34]	0.1702	[1.25]
PSTKRQH (redeemable preferred/preference stock)	0.1967	[1.19]	0.1901	[1.46]	0.1895	[1.40]
TEQQH (total stockholders equity)	0.1993	[1.21]	0.1957	[1.50]	0.1939	[1.43]
PPENTQH (total (net) property, plant, and equipment)	0.2045	[1.26]	0.1994	[1.50]	0.1852	[1.34]
LTQH (total liabilities)	0.1989	[1.22]	0.1929	[1.45]	0.1805	[1.30]
PSTKQH (total preferred/preference stock (capital))	0.1855	[1.12]	0.1778	[1.32]	0.1609	[1.14]
CEQQH (total common/ordinary equity)	0.1797	[1.09]	0.1744	[1.30]	0.1547	[1.11]
AOQH (total other assets)	0.2142	[1.39]	0.2341*	[1.85]	0.2252*	[1.71]
DLTTQH (total long-term debt)	0.2347	[1.62]	0.2617**	[2.12]	0.2838**	[2.21]
LOQH (total other liabilities)	0.2387*	[1.65]	0.2685**	[2.13]	0.3015**	[2.30]
ACOQH (total other current assets)	0.2664*	[1.77]	0.2920**	[2.27]	0.3313**	[2.48]
CHEQH (cash and short-term investments)	0.2622*	[1.72]	0.3386**	[2.46]	0.4658***	[3.35]
LCOQH (total other current liabilities)	0.2777*	[1.82]	0.3449**	[2.49]	0.4786***	[3.43]
APQH (accounts payable)	0.2660*	[1.74]	0.3407**	[2.51]	0.4863***	[3.59]
DVPQH (preferred/preference dividends)	0.2479	[1.62]	0.3261**	[2.40]	0.4679***	[3.46]
SALEQH (sales/turnover (net))	0.3711**	[2.51]	0.4474***	[3.49]	0.5579***	[4.27]
XIDOQH (extraordinary items and discontinued operations)	0.3427**	[2.33]	0.4293***	[3.39]	0.5294***	[4.09]
IBQH (income before extraordinary items)	0.5926***	[4.03]	0.7419***	[6.04]	0.7530***	[5.87]
IBADJQH (income before extraordinary items, adjusted for common stock equivalents)	0.6329***	[4.24]	0.7793***	[6.26]	0.7825***	[6.02]
NIQH (net income (loss))	0.6263***	[4.24]	0.7643***	[6.18]	0.7613***	[5.90]
IBCOMQH (income before extraordinary items, available for common)	0.6114***	[4.21]	0.7445***	[6.08]	0.7394***	[5.79]
PIQH (pretax income)	0.6551***	[4.49]	0.7815***	[6.43]	0.7733***	[6.10]
TXTQH (total income taxes)	0.6058***	[4.10]	0.7354***	[5.95]	0.7356***	[5.70]
NOPIQH (nonoperating income (expense))	0.6329***	[4.29]	0.7627***	[6.37]	0.7258***	[5.82]
DOQH (discontinued operations)	0.6463***	[4.44]	0.7802***	[6.55]	0.7495***	[6.04]
DVQH (cash dividends)	0.4814***	[3.19]	0.6232***	[5.11]	0.6133***	[4.83]
Panel B: Signal deletions (sequentially dropped variables)						
None (signal with all variables)	0.4814***	[3.19]	0.6232***	[5.11]	0.6133***	[4.83]
ATQH (total assets)	0.4877***	[3.23]	0.6287***	[5.13]	0.6190***	[4.86]
SEQQH (total stockholders equity)	0.4634***	[3.09]	0.6040***	[4.92]	0.5907***	[4.63]
ICAPTQH (total invested capital)	0.4523***	[3.01]	0.5876***	[4.77]	0.5701***	[4.45]
PSTKRQH (redeemable preferred/preference stock)	0.4338***	[2.90]	0.5707***	[4.65]	0.5516***	[4.32]
TEQQH (total stockholders equity)	0.4222***	[2.82]	0.5593***	[4.57]	0.5446***	[4.28]
PPENTQH (total (net) property, plant, and equipment)	0.4267***	[2.86]	0.5604***	[4.58]	0.5469***	[4.29]
LTQH (total liabilities)	0.3619**	[2.35]	0.5094***	[4.23]	0.4851***	[3.86]
PSTKQH (total preferred/preference stock (capital))	0.3785**	[2.44]	0.5252***	[4.31]	0.4974***	[3.92]
CEQQH (total common/ordinary equity)	0.4058***	[2.69]	0.5429***	[4.54]	0.5202***	[4.19]
AOQH (total other assets)	0.3963***	[2.85]	0.5381***	[4.47]	0.5721***	[4.75]
DLTTQH (total long-term debt)	0.3558**	[2.37]	0.4874***	[4.01]	0.4561***	[3.69]
LOQH (total other liabilities)	0.4135***	[2.72]	0.5192***	[4.23]	0.4408***	[3.49]
ACOQH (total other current assets)	0.3754**	[2.45]	0.4689***	[3.83]	0.3955***	[3.14]
CHEQH (cash and short-term investments)	0.3644**	[2.14]	0.3796***	[2.91]	0.2946**	[2.18]
LCOQH (total other current liabilities)	0.3357**	[1.98]	0.3504***	[2.68]	0.2821**	[2.08]
APQH (accounts payable)	0.3380**	[2.01]	0.3507***	[2.69]	0.2951**	[2.17]
DVPQH (preferred/preference dividends)	0.3422**	[2.01]	0.3566***	[2.71]	0.3005**	[2.20]
SALEQH (sales/turnover (net))	0.1898	[1.06]	0.2023	[1.56]	0.1097	[0.83]
XIDOQH (extraordinary items and discontinued operations)	0.1825	[1.03]	0.1971	[1.55]	0.1028	[0.78]
IBQH (income before extraordinary items)	0.1904	[1.08]	0.2149*	[1.68]	0.1210	[0.92]
IBADJQH (income before extraordinary items, adjusted for common stock equivalents)	0.1899	[1.07]	0.2126*	[1.65]	0.1121	[0.85]

(*continued on next page*)

Table 6 (*continued*)

Variables	Industry-adjusted return		Six-factor alpha		Eight-factor alpha	
	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]	Coefficient	[*t*-statistic]
NIQH (net income (loss))	0.1738	[0.98]	0.2034	[1.58]	0.1108	[0.83]
IBCOMQH (income before extraordinary items, available for common)	0.1427	[0.81]	0.1634	[1.27]	0.0839	[0.63]
PIQH (pretax income)	0.0067	[0.04]	−0.0033	[−0.02]	−0.0733	[−0.52]
TXTQH (total income taxes)	−0.0598	[−0.35]	−0.0865	[−0.64]	−0.0203	[−0.15]
NOPIQH (nonoperating income (expense))	−0.0607	[−0.35]	−0.0889	[−0.66]	−0.0413	[−0.30]
DOQH (discontinued operations)	−0.0589	[−0.34]	−0.0909	[−0.67]	−0.0363	[−0.26]
DVQH (cash dividends)	−0.1172	[−0.70]	−0.1559	[−1.20]	−0.0718	[−0.54]
Panel C: Balance sheet, income and cash flow statement items						
All balance sheet items	0.2643*	[1.73]	0.3396**	[2.50]	0.4852***	[3.59]
All income and cash flow statement items	0.3380**	[2.01]	0.3507***	[2.69]	0.2951**	[2.17]

mispricing signal is unlikely to capture omitted risk factors tied to market capitalization. This finding buttresses our earlier argument that the Berk (1995) critique does not apply here.

The two Panel C specifications separately analyze the efficacy of the balance sheet items compared with the income and cash flow items. Using the 16 balance sheet items generates about as much performance as the income statement and cash flow items. In sum, Table 6 shows that more parsimonious ways of estimating mispricing exist, but the precise specification of that parsimony is hard to flesh out. Also, any more parsimonious approach to identifying fair value cannot consist of just a handful of items, and it cannot consist of items solely from the income statement, balance sheet, or cash flow statement. To further illustrate this point, a more discretionary signal that uses only three accounting items (book equity, net income, and sales) generates mostly statistically insignificant alphas that are about 50% smaller compared with those of our more agnostic signal derived from 28 accounting items.

We also implement a mechanical stepwise procedure that adds and deletes regressors, one at a time, from the valuation regression model for that month. Using a 15% significance level, the procedure settles on a specification in which each regressor in the model is individually significant and each potential regressor omitted is insignificant when added to the model. Averaged across sample months, the procedure selects about 21 of the 28 variables (minimum 15, maximum 27). The risk-adjusted performance is, however, similar to that from the signal that uses the same 28 variables in each month's valuation regression.

7. Conclusion

Regression-based fitting of the latest available accounting data to stock values leads to a mispricing measure constructed from regression residuals. This hedonic approach to fundamental analysis is essentially saying that firms should be valued as a collection of accounting attributes from their most recent accounting statements. With our approach, future returns play no role for the weighting of the accounting items, yet ranking firms based on their residual-implied percentage mispricing predicts returns in the subsequent month and up to three years in the future. The results are not related to the most commonly known predictors of the cross-section of expected returns. Abnormal return spreads based on mispricing metrics formed from accounting data range from 4% to 10% per year. Thus, market prices do not fully reflect accounting data. Rudimentary statistical analysis of the most commonly reported accounting information leads to risk-adjusted returns of a magnitude comparable to those earned by value and momentum strategies.

Our industry-adjusted returns and alphas tend to be negative across mispricing quintiles and are not significantly positive for the underpriced firms. This is due to the industry adjustment as the equal-weighted industry portfolios have much more weight on low priced stocks than either the mispricing portfolios or the benchmark factor portfolios.[30]

One could investigate other potentially valuable information with the type of statistical analysis undertaken here. The other information could include changes in the same item in more than one consecutive accounting statement, analyst forecasts, or corporate actions. One could even combine the current accounting information with the information in past price movements. An investigation of these issues is beyond the scope of this paper. Our task here is to apply minimal discretion in estimating a peer-implied intrinsic value to examine if a reasonably agnostic form of fundamental analysis works. It seems to work very well.

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 and convergence to fair value. We have, however, presented evidence supporting the claim that the abnormal profits earned from our version of fundamental analysis are not due to an omitted risk factor. We focus only on returns, adjusted for risk factors, instead of more direct measures of convergence, because measuring convergence from returns is a more conservative approach. Our estimate of fair value exhibits regression toward its mean (of true fair value)

[30] With no cutoffs for share price, the low-priced stocks in the French data library's equally weighted industry sample load more negatively on the momentum and RMW earnings factors (boosting equally weighted industry portfolio alphas significantly, perhaps artificially). Unlike most anomaly studies, our raw sample does not have this alpha issue because of its $5 share price filter. With a customized industry portfolio adjustment restricted to our sample selection or with no industry adjustment at all, there is no tendency toward negative alphas across quintiles.

over time, like most other estimates. Hence, direct measure of the dynamics of the distance between estimated fair value and prices leads to stronger convergence estimates than examining returns alone. 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.

Because we focus indiscriminately on the most widely available accounting items, and because of their high degree of collinearity, successfully identifying which accounting variables are best for determining fair value is difficult. Addressing this issue more precisely with a try-all-specifications approach is blatant data snooping. Because the accounting data seem to have an underlying factor structure, many fewer accounting variables could do as well as, or improve upon, the strategies derived here. We leave that, and improvements in the fair value estimation approach, to future research.

Our paper is not another anomaly paper because our approach differs from the approaches taken by the papers in the anomalies literature. The selection of accounting items is intended to be universal, except that coverage and statistical power require limitations on the number of accounting items. Therefore, as a practical compromise, our model employs the most common reported accounting items across firms. The absence of discretion here distinguishes our paper from predecessors that study market efficiency and represents its unique contribution to that literature.

Obvious steps for future research would extend this analysis to global financial markets and to address whether firms are aware of how their share prices deviate from fair value. With respect to the first extension of this paper's research, Bartram and Grinblatt (2017) have already commenced such analysis, finding that the mispricing signal works particularly well in regions with frictions that deter arbitrageurs. With respect to the second, preliminary results show that the mispricing signal predicts the repurchasing and issuing of shares by companies over the subsequent three to 12 months. These new findings further confirm that markets do not always set prices equal to fundamental values.

Appendix A. Discussion and proof of result in the introduction

For a given date, let X^* denote the $N \times K$ 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 instead of per share), preserving the linearity of valuation.[31] Thus, the accounting items that would be reported for an investment in the N firms represented by the $1 \times N$ vector w would be the $1 \times K$ vector wX^*. More generally, for N distinct investments given by the rows of the $N \times N$ matrix W, the accounting statements of the investments would be given by the rows of the $N \times K$ matrix WX^*. Thus, the N replicating investments must satisfy

$$WX^* = X^*, \tag{4}$$

and if the associated fair value estimates are further required to have average mispricing of zero, then W must satisfy

$$WX = X, \tag{5}$$

where the $N \times (K+1)$ matrix X is X^* augmented by a (first) column of 1s.[32] 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 1. A unique W of rank $K+1$ that is a function of X produces zero average mispricing and minimizes the mean-squared prediction error of any non-accounting attribute v of the targets. This W is the one given by the idempotent projection matrix that statisticians are familiar with from linear regression.

Proof. Project any variable y not spanned by X onto X, which decomposes

$$y = X\left(X^T X\right)^{-1} X^T y + \varepsilon, \tag{6}$$

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. Because 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

$$W = X\left(X^T X\right)^{-1} X^T \tag{7}$$

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

$$P = Wv = X\left(X^T X\right)^{-1} X^T v. \tag{8}$$

Appendix B. Variable definitions

This Appendix contains the variable name (or mnemonic), the description (or construction) of the data item, and the source (database) in parentheses. The Center for Research in Security Prices (CRSP) and Compustat Point-in-time (PIT) are from the Wharton Research

[31] For example, the revenue of an investment that buys up 100% 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 exchange-traded funds instead of 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.

[32] This means that the N eigenvalues of W consist of K+1 1s and N−K−1 0s. 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 1s, 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.

Data Services (WRDS). The Kenneth French website can be found at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

ATQH: assets, total, quarterly (Compustat).
DVPQH: dividends, preferred/preference, quarterly (Compustat).
SALEQH: sales/turnover (net), quarterly (Compustat).
SEQQH: stockholders equity, total, quarterly (Compustat).
IBQH: income before extraordinary items, quarterly (Compustat).
NIQH: net income (loss), quarterly (Compustat).
XIDOQH: extraordinary items and discontinued operations, quarterly (Compustat).
IBADJQH: income before extraordinary items, adjusted for common stock equivalents, quarterly (Compustat).
IBCOMQH: income before extraordinary items, available for common, quarterly (Compustat).
ICAPTQH: invested capital, total, quarterly (Compustat).
TEQQH: stockholders equity, total, quarterly (Compustat).
PSTKRQH: preferred/preference stock, redeemable, quarterly (Compustat).
PPENTQH: property, plant, and equipment, total (net), quarterly (Compustat).
CEQQH: common/ordinary equity, total, quarterly (Compustat).
PSTKQH: preferred/preference stock (capital), total, quarterly (Compustat).
DLTTQH: long-term debt, total, quarterly (Compustat).
PIQH: pretax income, quarterly (Compustat).
TXTQH: income taxes, total, quarterly (Compustat).
NOPIQH: non-operating income (expense), quarterly (Compustat).
AOQH: assets, other, total, quarterly (Compustat).
LTQH: liabilities, total, quarterly (Compustat).
DOQH: discontinued operations, quarterly (Compustat).
LOQH: liabilities, other, total, quarterly (Compustat).
CHEQH: cash and short-term investments, quarterly (Compustat).
ACOQH: current assets, other, total, quarterly (Compustat).
DVQH: cash dividends (cash flow), quarterly (Compustat).
LCOQH: current liabilities, other, total, quarterly (Compustat).
APQH: accounts payable, quarterly (Compustat).
Share price: stock price (in dollars 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 (Kenneth French website).
Industry portfolios: monthly returns on 38 industry portfolios (Kenneth French website).
Mkt_RF: monthly market index return net of risk-free rate (Kenneth French website).
SMB: monthly small minus big portfolio return (Kenneth French website).
HML: monthly high minus low portfolio return (Kenneth French website).
Mom: monthly momentum portfolio return (Kenneth French website).
ST_Rev: monthly short-term reversal portfolio return (Kenneth French website).
LT_Rev: monthly long-term reversal portfolio return (Kenneth French website).
CMA: monthly investment factor (Kenneth French website).
RMW: monthly profitability factor (Kenneth French website).
SUE: quarterly unexpected earnings surprise based on a rolling seasonal random walk model (Livnat and Mendenhall, 2006, p. 185).
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 (ATQH) less cash and short-term investments (CHEQH). Operating liabilities is calculated as total assets (ATQH) less total debt (DLCQH and DLTTQH) less book value of total common and preferred equity (CEQQH and PSTKQH) less minority interest (MIBTQH) (Richardson, Sloan, Soliman, and Tuna 2001, p. 22).
Gross profitability: (revenue (SALEQH) – cost of goods sold (COGSQH))/total assets (ATQH) (Novy-Marx, 2013).
Market capitalization: stock market capitalization of common stock, calculated as product of share price (PRC) * number of shares outstanding (SHROUT).
Book/market: (book equity (CEQQH) + deferred taxes balance sheet (TXDITCQH))/market capitalization
Mispricing signal (M): −1 * residual/market capitalization.
Short-term reversal: return in prior month.
Momentum: return in prior year excluding prior month.
Long-term reversal: return in prior five years excluding prior year.
Scaled NOA: scaled net operating assets (NOA) (Hirshleifer, Hou, Teoh, and Zhang, 2004).
Share issuance: share issuance (Daniel and Titman, 2006).
Asset growth: asset growth (Cooper, Gulen, and Schill, 2008).
Capital investment: abnormal capital investment (Titman, Wei, and Xie, 2004).
Investment ratio: investment ratio (Lyandres, Sun, and Zhang, 2008).
External financing: external financing (Bradshaw, Richardson, and Sloan, 2006).
Z-score: Z-score (Ferguson and Shockley, 2003).
Leverage: leverage ratio (Ferguson and Shockley, 2003).
Illiquidity: illiquidity measure (Amihud, 2002).
Earnings yield: earnings to price (Penman, Richardson, Riggoni, and Tuna, 2014).
Dividend/price: dividends to price (Fama and French, 1992).
Cash flow/price: cash flow to price (Hou, Karolyi, and Kho, 2011).

V/P: value to price (Lee, Myers, and Swaminathan, 1999).

References

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

Abarbanell, J.S., Bushee, B.J., 1998. Abnormal returns to a fundamental analysis strategy. Accounting Review 73, 19–45.

Acharya, V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375–410.

Amihud, Y., 2002. Illiquidity and stock returns: cross section and time series effects. Journal of Financial Markets 5, 31–56.

Aretz, K., Bartram, S.M., Pope, P.F., 2010. Macroeconomic risks and characteristic-based factor models. Journal of Banking and Finance 34, 1383–1399.

Avramov, D., Chordia, T., Jostova, G., Philipov, A., 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., Bartov, E., 1996. How naive is the stock market's use of earnings information? Journal of Accounting and Economics 21, 319–337.

Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159–178.

Bartram, S.M., Grinblatt, M., 2017. Global Market Inefficiencies. University of Warwick, Coventry, UK Unpublished Working Paper.

Beneish, M.D., Lee, C.M.C., Nichols, D.C., 2015. In short supply: shortsellers and stock returns. Journal of Accounting and Economics 60, 33–57.

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

Berk, J., van Binsbergen, J.H., 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics 118, 1–20.

Bernard, V.L., Thomas, J., 1989. Post-earnings announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1–36.

Bernard, V.L., Thomas, J., 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.

Bessembinder, H., Cooper, M.J., Zhang, F., 2015. Characteristic-Based Expected Returns and Corporate Events. University of Utah, Salt Lake City, UT Unpublished Working Paper.

Bhojraj, S., Lee, C.M.C., 2002. Who is my peer? a valuation-based approach to the selection of comparable firms. Journal of Accounting Research 40, 407–439.

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

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899–2939.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chen, H.L., Jegadeesh, N., Wermers, R., 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, 343–368.

Cooper, I.A., Lambertides, N., 2014. Is There a Limit to the Accuracy of Equity Valuation Using Multiples?. London Business School, London, UK Unpublished working paper.

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

Daniel, K.D., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Daniel, K.D., Hirshleifer, D., Subrahmanyam, A., 1998. A theory of overconfidence, self-attribution, and security market under- and over-reactions. Journal of Finance 53, 1839–1885.

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

DeBondt, W., Thaler, R., 1985. Does the stock market overreact? Journal of Finance 40, 793–805.

Dechow, P.M., Hutton, A.P., Sloan, R.G., 1999. An empirical assessment of the residual income valuation model. Journal of Accounting and Economics 26, 1–34.

Dichev, I.D., 1998. Is the risk of bankruptcy a systematic risk? Journal of Finance 53, 1131–1147.

Dong, M., Hirshleifer, D., Richardson, S., Teoh, S.H., 2006. Does investor misvaluation drive the takeover market? Journal of Finance 61, 725–762.

Edmans, A., Goldstein, I., Jiang, W., 2012. The real effects of financial markets: the impact of prices on takeovers. Journal of Finance 67, 933–971.

Fama, E.F., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.

Fama, E.F., French, K.R., 1992. The cross section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

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

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

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

Fama, E.F., French, K.R., 2014. A Five-Factor Asset Pricing Model. University of Chicago, Chicago, IL Unpublished Working Paper.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.

Ferguson, M.F., Shockley, R.L., 2003. Equilibrium "anomalies. Journal of Finance 58, 2549–2580.

Foster, F.D., Smith, T., Whaley, R.E., 1997. Assessing goodness-of-fit of asset pricing models: the distribution of the maximal R^2. Journal of Finance 52, 591–607.

Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. Accounting Review 59, 574–603.

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

Frazzini, A., Pedersen, L.H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.

Gatev, E., Goetzmann, W.N., Rouwenhorst, K.G., 2006. Pairs trading: performance of a relative-value arbitrage rule. Review of Financial Studies 19, 797–827.

Gerakos, J., Linnainmaa, J., 2017. Decomposing value, *Review of Financial Studies*, (forthcoming).

Green, J., Hand, J., Zhang, F., 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, 413–442.

Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. Journal of Financial Economics 78, 311–339.

Grinblatt, M., Jostova, G., Petrasek, L., Philipov, A., 2017. Style and Skill: Hedge Funds, Mutual Funds, and Momentum. University of California, Los Angeles, CA Unpublished Working Paper.

Grinblatt, M., Titman, S., 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 393–416.

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

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

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

Habib, M., Ljungqvist, A., 2005. Firm value and managerial incentives: a stochastic frontier approach. Journal of Business 78, 2053–2094.

Harvey, C., Liu, Y., Zhu, H., 2016. … and the cross section of expected returns. Review of Financial Studies 29, 5–68.

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

Hirshleifer, D., Jiang, D., 2010. A financing-based misvaluation factor and the cross section of expected returns. Review of Financial Studies 23, 3401–3436.

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

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

Hou, K., Karolyi, G.A., Kho, B.-C., 2011. What factors drive global stock returns? Review of Financial Studies 24, 2527–2574.

Ikenberry, D., Lakonishok, J., Vermaelen, T., 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., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: an evaluation of alternative explanations. Journal of Finance 56, 699–720.

Jiang, H., Verbeek, M., Wang, Y., 2014. Information content when mutual funds deviate from benchmarks. Management Science 60, 2038–2053.

Jones, C.P., Litzenberger, R.H., 1970. Quarterly earnings reports and intermediate stock price trends. Journal of Finance 25, 143–148.

Joy, O.M., Litzenberger, R.H., McEnally, R.W., 1977. The adjustment of stock prices to announcements of unanticipated changes in quarterly earnings. Journal of Accounting Research 15, 207–225.

Kogan, L., Tian, M., 2013. Firm Characteristics and Empirical Factor Models: A Data-Mining Experiment. Massachusetts Institute of Technology, Cambridge, MA Unpublished Working Paper.

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

Lee, C.M.C., Myers, J., Swaminathan, B., 1999. What is the intrinsic value of the Dow? Journal of Finance 54, 1693–1741.

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

Liu, J., Nissim, D., Thomas, J., 2002. Equity valuation using multiples. Journal of Accounting Research 40, 135–172.

Livnat, J., Mendenhall, R., 2006. Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts. Journal of Accounting Research 44, 177–205.

Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–467.

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

Lyandres, E., Sun, L., Zhang, L., 2008. The new issues puzzle: Testing the investment-based explanation. Review of Financial Studies 21, 2827–2855.

Manaster, S., Rendleman, R.J., 1982. Option prices as predictors of equilibrium stock prices. Journal of Finance 37, 1043–1057.

McLean, D., Pontiff, J., 2016. Does academic research destroy stock return predictability? Journal of Finance 71, 5–32.

Michaely, R., Thaler, H., Womack, L., 1995. Price reactions to dividend initiations and omissions: overreaction or drift? Journal of Finance 50, 573–608.

Mitchell, M.L., Stafford, E., 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, 133–170.

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

Ohlson, J.A., 1990. A synthesis of security valuation theory and the role of dividends, cash flows, and earnings. Contemporary Accounting Research 6, 648–676.

Ohlson, J.A., 1991. The theory of value and earnings, and an introduction to the Ball-Brown analysis. Contemporary Accounting Research 7, 1–19.

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

Ohlson, J.A., Kim, S., 2015. Linear valuation without OLS: the Theil-Sen estimation approach. Review of Accounting Studies 20, 395–435.

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

Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

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

Perold, A.F., Tierney, B., 1997. Numeric Investors L.P. Case study 9-298-012. Harvard Business School Publishing, Boston, MA.

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.

Piotroski, J.D., So, E.C., 2012. Identifying expectation errors in value/glamour strategies: a fundamental analysis approach. Review of Financial Studies 25, 2841–2875.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. Journal of Finance 93, 921–945.

Rendleman, R.J., Jones, C.P., Latane, H.A., 1982. Empirical anomalies based on unexpected earnings and the importance of risk adjustments. Journal of Financial Economics 10, 269–287.

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

Richardson, S.A., Sloan, R.G., Soliman, M., Tuna, I., 2001. Information in accruals about the quality of earnings. London Business School, London, UK Unpublished Working Paper.

Ross, S.A., 1978. A simple approach to the valuation of risky streams. Journal of Business 51, 453–475.

Schulmeister, S., 2009. Profitability of technical stock trading: has it moved from daily to intraday data? Review of Financial Economics 18, 190–201.

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

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association 63, 1379–1389.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52, 327–340.

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

Stambaugh, R.F., Yu, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

Sullivan, R., Timmerman, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. Journal of Finance 54, 1647–1691.

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

Theil, H., 1950. A rank-invariant method of linear and polynomial regression analysis. Nederlandse Akademie Wetenchappen, Series A 53, 386–392.

Titman, S., Wei, K.C.J., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.

Wermers, R., 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style, transaction costs, and expenses. Journal of Finance 4, 1655–1695.