Which stock return predictors reflect mispricing? *

Jonas Frey†

November 22, 2023

Abstract

A large number of stock characteristics have been found to predict the cross-section of returns. Return predictability can be driven by risk or mispricing, and the nature of most return predictors remains an open question. I use analysts' earnings forecasts to determine if a return predictor is linked to mispricing. I find that at least 40% of return predictors from a dataset of 172 significant predictors are related to mispricing, including the momentum predictor from the Carhart four-factor and the profitability and investment predictors from the Fama–French five-factor model. I further study whether the mispricing predictors' abnormal returns capture the divergence of prices from the fundamental value (build-up predictors) or their convergence back to the fundamental value (resolution predictors). Build-up predictors are less common than resolution predictors but they do exist, implying that trading on certain return predictors can exacerbate rather than eliminate mispricing. Momentum is related both to the build-up and the resolution of mispricing.

*I would like to thank Pedro Bordalo, Michael Halling (Discussant), Cameron Peng and Mungo Wilson, participants of the EFA 2023 Doctoral Tutorial, as well as seminar participants at Oxford and the Oxford/London Behavioural Finance Group for helpful comments. The author acknowledges the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work.

†Saïd Business School, University of Oxford & University of Münster, jonas.frey.dphil2017@said.oxford.edu

1 Introduction

Over the past decades, hundreds of cross-sectional stock return predictors have been discovered. For instance, small-cap stocks offer higher returns than large-cap stocks and stocks with a high book-to-market ratio outperform those with a low book-to-market ratio (Fama and French, 1992). The vast majority of these predictors were discovered empirically rather than being derived from theory. Therefore, for most predictors, there is no consensus on the channel through which they drive returns.

Given that the fundamental value of a stock is its discounted stream of dividends, stock returns over a given period can be driven by two possible components: the discount factor applied to future dividends and changes in dividend expectations. While early studies suggested that returns are almost exclusively driven by changes in required returns (i.e., discount factors), more recent studies suggest that changes in dividend expectations play a significant role (L. Chen et al., 2013; De La O and Myers, 2021; Pruitt, 2023). Moreover, risk-based models have struggled to fully explain the volatility of asset returns across a wide range of asset classes (Giglio and Kelly, 2018). At the same time, previous research using data on expectations demonstrates that expectations are biased and that these biases are linked to return predictability (e.g., Bordalo et al., 2019, 2023; La Porta, 1996). Therefore, it is likely that at least some of the many known stock return predictors are not proxies for systematic risk but for mispricing.

In this paper, I develop a method to test if a given return predictor is associated with mispricing driven by biased expectations, as opposed to being exclusively driven by risk or mispricing from other sources, such as market frictions. If a return predictor is driven by risk, it is linked to returns through the discount factor. Specifically, the more priced risk a stock is exposed to, the more investors will discount its future dividends, leading to a higher return. In contrast, changes in dividend expectations are unrelated to risk. Moreover, under (full information) rational expectations, they should be unpredictable because all available information is already incorporated optimally into the current forecast and can thus not predict future changes in this forecast. Since dividend expectations drive prices, biased expectations imply mispricing. Therefore, a predictor that can predict changes in dividend expectations in addition to returns is linked to mispricing.

For my test, I use the earnings forecast of professional analysts as a proxy for the market's dividend expectations[1]. Given the argument above, I can say that a return predictor is a mispricing predictor (henceforth also called a "mispricing") if it also predicts more positive earnings forecast revisions for the stocks for which it predicts more positive returns. In principle, a link between a return predictor and biased expectations could stem from behavioural biases or information processing frictions. However, since the predictors use simple portfolio sorts based on widely available data, it is unlikely that frictions prevent market participants from incorporating the predictor's information into the price.

My test provides four crucial insights by identifying the return predictors that are associated with biased expectations, which are a subset of all return predictors linked to mispricing. First, mispricing of any kind is evidence of market inefficiency, whereas return predictability from risk is not. Thus, my test provides a lower bound on the share of return predictability stemming from mispricing. Understanding how much mispricing exists in the market is crucial because prices determine capital allocation and can thus affect the real economy.

Second, mispricing-driven return predictors should only be used in descriptive asset pricing models and should not be used normatively, such as to calculate risk-adjusted returns.

Third, mispricing predictors whose returns are at least partially driven by biased expectations suggest profitable trading strategies, which is not the case for risk or mispricing from frictions. Therefore, the fact that my test only captures mispricing related to biased expectations and not friction-driven mispricing can be seen as an advantage.

Finally, mispricing from biased expectations can be eliminated through trading once arbitrageurs become aware of it, whereas return predictability from risk or market frictions should be persistent. Crucially, as long as arbitrageurs are uncertain about whether a return

[1]A large amount of previous work shows that analyst expectations are linked to returns and can be used as a proxy for the market's expectations. See Adam and Nagel (2023) for a recent review of the literature.

predictor reflects mispricing or risk, they do not know if it offers abnormal risk-adjusted returns, and they will be reluctant to trade based on it. Therefore, research on the nature of return predictors can lead to greater market efficiency by informing the trading decisions of arbitrageurs.

In summary, given the key differences between risk-driven and the various types of mispricing-driven return predictors, we arguably learn very little from knowing about the existence of a return predictor if we do not understand its nature. Establishing the nature of a predictor is challenging because predictor discovery is mainly driven by empirical research, and it is difficult to establish a predictor's nature from the return distribution without an economic model (Kozak et al., 2018). Therefore, considerable disagreement about the nature of most predictors remains in the literature (see Holcblat et al., 2022, for a review) and arguably the CAPM market factor (Lintner, 1965; Sharpe, 1964) is the only predictor that is universally accepted to represent risk and not mispricing. Thus, it is crucial to develop portable empirical tests that can help understand if a given return predictor is driven by risk or mispricing.

The remainder of the paper proceeds as follows: In Section 2, I describe my data and methodology. I use a dataset created by A. Y. Chen and Zimmermann (2021) that contains monthly stock returns as well as data on 212 characteristics, which showed significant evidence of return predictability in the paper that introduced them, (predictors) as well as 113 characteristics that lacked such evidence (placebos). I merge this data with analysts' earnings forecasts from IBES.

In Section 3, I first show that 172 out of the 212 predictors offer significant return predictability if they are extended beyond their original sample period using data up to the end of 2022. Applying my mispricing test to these 172 predictors, I find that around 40% of them also predict earnings forecast revisions and are hence linked to mispricing. Importantly, my test cannot rule out that a predictor is also associated with risk in addition to being linked to mispricing.

The mispricing predictors include the momentum factor from the Carhart (1997) four-factor model as well as the profitability and the investment factor from the Fama and French (2015) five-factor model. Moreover, 4 out of the 14 factors that survive the stepwise factor selection procedure by Feng et al. (2020) for which I have data are classified as mispricing, highlighting that there are robust factors among the mispricings. The types of predictors among which mispricing is particularly common are those related to external financing, volatility, earnings forecasts, and earnings growth, as well as the predictors that capture a lead-lag relationship.

Papers that discover a new predictor often offer guidance on whether the authors believe it represents risk or mispricing, even though they usually do not have a formal test or model. However, recent evidence suggests that the interpretation in the original paper can be unreliable (A. Y. Chen et al., 2022). In line with that, a comparison between my classification and the interpretation of the original authors suggests that at least 40% of the return predictors that were originally interpreted as risk factors are linked to mispricing. Therefore, either formal models or empirical tests, like the one I propose in this paper, are required to interpret the nature of return predictors.

A potential concern for the validity of my results is that some of the return predictors in my sample may be spurious and offer no real out-of-sample return predictability (Chordia et al., 2020; Harvey et al., 2016; Hou et al., 2020). My test could potentially classify such predictors as mispricing because stocks that are ex-post selected to have high (low) returns plausibly also have more positive (negative) news which in turn implies more positive (negative) earnings forecast revisions. This can lead to spurious predictors being classified as mispricing (Engelberg et al., 2018). I conduct two analyses to assess how likely this is to occur. First, I apply my test to the 68 predictors that had insignificant evidence of return predictability both in the original paper that discovered them and in my sample. I find that in this dataset, around 26% of the predictors are classified as mispricing, which is less than in the sample of significant predictors. While some insignificant predictors are classified as mispricing, it is also plausible that there are true return predictors that just lack statistical significance in this sample. Therefore, I perform a second test in which I generate simulated spurious predictor portfolios by randomly sorting stocks into long and short portfolios and retaining those portfolios that offer significant and sizeable long-minus-short returns by chance. I find that, in my preferred regression specification, not a single spurious predictor is associated with mispricing. This suggests that falsely classifying spurious predictors as mispricing is unlikely to be relevant in practice.

The abnormal returns associated with each mispricing can either reflect the convergences of prices back to fundamental values (resolution of mispricing) or their divergence from fundamental values (build-up of mispricing). In Section 4, I study which mispricings capture the build-up and the resolution of mispricing respectively. In a recent paper, van Binsbergen et al. (2023) perform this classification using a method that relies on calculating the fundamental value of stocks based on realised dividends and an assumption about the correct stochastic discount factor (SDF). They find that although most predictors capture the resolution of mispricing, a relevant share also captures its build-up. They also highlight the practical importance of this distinction: traders who exploit build-up predictors exacerbate mispricing rather than reduce it.

My classification method is based on the predictability of forecast errors, an orthogonal approach that does not require a stance on the correct SDF. For resolution predictors, stocks that predictably earn higher returns are initially underpriced relative to those earning lower returns. The subsequent higher returns then correct this mispricing. Hence, earnings expectations should also initially be significantly more pessimistic for stocks with predicted high returns relative to those with low returns. I find that around 57% of the mispricings can be classified as resolution. For build-up predictors, all stocks are initially priced correctly, and the subsequent higher returns of some stocks cause them to be overpriced relative to those that earned lower returns[2]. This implies that, after the period of excess returns, earnings

[2]It is also possible that the stocks earning higher returns are already overpriced initially, and become more overpriced over time

expectations are too positive for stocks that earned higher returns relative to those with lower returns. I find that around 24% of mispricing predictors can be classified as build-up. Importantly, it is possible that the predictably higher returns of initially underpriced stocks not only correct the underpricing but overshoot and cause the stocks to be overpriced. Thus, a return predictor can capture both the resolution of initial mispricing and the build-up of new mispricing. I find that this is indeed the case for most mispricings related to momentum, suggesting that momentum both reflects an initial underreaction and a delayed overreaction to good news. This result extends a recent finding by Cuevas Rodriguez et al. (2023), who show that momentum is linked to underreaction in analysts' expectations.

I am not the first to provide a method to test if a return predictor is associated with mispricing or risk. Pukthuanthong et al. (2019) develop a set of conditions that categorise a predictor as a risk factor if it is related to the covariance matrix of returns, priced and has a reasonable Sharpe ratio[3]. Gafka et al. (2021) compare predictors' ability to forecast returns on days when important information is released onto the market (announcement days) and the remaining days. They argue that if the return predictability of a given predictor is concentrated on announcement days, it is likely associated with risk. Closest to my work is a test by Holcblat et al. (2022) that classifies a predictor as mispricing if the hypothesis that every risk-averse individual would prefer to invest in its long portfolio rather than its short portfolio cannot be rejected.

My test has several unique features relative to previous work. It is the first test to specifically identify mispricing from biased expectations, whereas previous tests separate return predictability that can be explained by risk from all other types of return predictability. Hence, it allows for a more precise understanding of mispricing in the market. As discussed above, this distinction is meaningful because, unlike mispricing stemming from market frictions, mispricing from expectational biases suggests profitable trading strategies and can be eliminated through trading. Furthermore, the fact that my test is the only one to directly

[3]An older paper by Charoenrook and Conrad (2005) follows a similar approach but only derives a necessary but not sufficient condition.

detect mispricing as opposed to classifying a predictor as mispricing if its returns cannot be explained by risk also makes it the only test where a lack of statistical power will cause an underdetection rather than an overdetection mispricing. This makes it more suited to applications where failing to detect some mispricing predictors is less concerning than falsely identifying mispricings, such as deciding on a trading strategy. Notably, all existing tests share the drawback that they can only provide evidence that a certain predictor is linked to mispricing (risk) but cannot rule out that it is also linked to risk (mispricing).

In addition to issues with interpretability, the sheer number of return predictors poses a challenge to asset pricing research, as it is not tractable to work with models that include hundreds of predictors. An important first step to reduce the number of predictors is to remove those that are spurious (i.e., offer no out-of-sample return predictability). Several recent papers have addressed this issue (Chordia et al., 2020; Giglio et al., 2021; Harvey et al., 2016; Hou et al., 2020). Going beyond validity, it has been shown that several predictors capture the same source of underlying return predictability making it unnecessary to include each of them in a model. As a response, methods to select the relevant return predictors or combine different return predictors were developed in order to shrink the number of relevant factors (Feng et al., 2020; Harvey and Y. Liu, 2021; He et al., 2022; Jensen et al., 2023).

I add to this literature in two ways. First, I show that a predictor that is classified as mispricing by my test is unlikely to be spurious, implying that the test not only provides evidence on the nature but also the validity of a predictor. Second, I show that the return predictability of around 40% of the predictors is associated with biased expectations. This suggests that a new return predictor capturing the biased expectations could potentially subsume the return predictability of a large set of existing predictors. However, developing such a predictor is beyond the scope of this paper[4].

Finally, my paper contributes to the literature that shows links between expectation data

[4]Developing an expectation bias return predictor is challenging because it needs to predict returns and, hence, cannot be based on ex-post forecast errors. Moreover, build-up and resolution predictors predict different levels of bias at different points in the return cycle.

and asset returns (see Adam and Nagel (2023) for a review) by showing that this data can not only be used directly to predict returns but also to understand the nature of return predictor characteristics. Some previous work in this literature has linked individual return predictors (Bouchaud et al., 2019; Bradshaw et al., 2006; Cuevas Rodriguez et al., 2023; Jackson and Johnson, 2006; La Porta et al., 1997) or small sets of predictors (Ben-Rephael et al., 2021; Bordalo et al., 2023; Grinblatt et al., 2018) to biased expectations[5]. Closest to my paper is a strand of the literature showing that analysts are more pessimistic (optimistic) about stocks with predictably higher (lower) returns. (van Binsbergen et al., 2022; Engelberg et al., 2018; Kozak et al., 2018). Notably, this approach implicitly assumes that all mispricings are resolution predictors since the opposite should be the case for build-up predictors[6]. I add to this literature by providing a test for the nature of a return predictor based on forecast revisions over the return period that is valid for both build-up and resolution predictors. Moreover, I provide results on the share of mispricings among a large set of predictor variables and an individual classification for each predictor. In contrast, previous research either studied only a few predictors or aggregated a large set of predictors into a single score, which does not allow conclusions about the nature of individual predictors. Finally, I provide a more detailed classification by further separating mispricing predictors into build-up and resolution predictors.

[5]A tangentially related literature also shows that analysts' stock return predictions (rather than firm earnings predictions) are related to return predictors (Engelberg et al., 2020; L. Guo et al., 2020).

[6]For a build-up predictor, there is no difference in expectations at portfolio formation. Depending on the source of the mispricing, analysts should be either too optimistic about stocks in the long portfolio or too pessimistic about stocks in the short portfolio at the end of the return period. In either case, average expectations should be more positive for stocks in the long portfolio than those in the short one.

2 Data & Methodology

2.1 Terminology

In this section, I describe my terminology and how it relates to other terms used in the literature. I use the term *return predictor* to describe any variable that can predict the cross-section of returns, regardless of the source of return predictability. A common alternative term used to describe the same concept in the literature is *factor*, which stems from the fact that return predictability is often studied in the context of factor models. Moreover, some researchers use the term *anomaly* to describe all variables that can predict returns, except the CAPM Beta.

I use the term *mispricing* for every return predictor that is not exclusively driven by risk. This includes return predictability due to behavioural biases or frictions. Other research sometimes also uses the terms *anomaly* or *characteristic* for these types of return predictors.

2.2 Motivation of the Empirical Specification

In this section, I motivate the empirical specification used to detect mispricings among a set of predictors potentially containing mispricings and risk factors. The test will not be able to detect mispricing from any source but only mispricing driven by biased expectations about a firm's future prospects. Derivations of the equations in this section can be found in Appendix A.

The most common way to establish that a characteristic can be used to predict stock returns is the following procedure: First, the cross-section of stocks is assigned to long and short portfolios at regular intervals, based on the level of the characteristic for each stock at the time of portfolio formation and the hypothesised relation between the characteristic and returns. For example, for the size predictor, the 10% of stocks with the lowest market capitalisation are assigned to the long portfolio and the 10% of stocks with the highest market capitalisation are assigned to the short portfolio. In the second step, the long-minus-short return is calculated for each month by subtracting the (potentially weighted) average return of stocks in the short portfolio from that of stocks in the long portfolio. A characteristic is then considered to be a return predictor if the average long-minus-short return across time is significantly larger than zero.

To better understand the source of return predictability, I decompose the long-minus-short return of a predictor characteristic into several components. I start with a stock level decomposition of the return, first proposed by Campbell (1991) based on the Campbell-Shiller decomposition (Campbell and Shiller, 1988a,b).

$$r_{i,t+1} = \mathbb{E}_t(r_{i,t+1}) + [\mathbb{E}_{t+1} - \mathbb{E}_t]\sum_{s=0}^{\infty}\rho_i^s g_{i,t+1+s} - [\mathbb{E}_{t+1} - \mathbb{E}_t]\sum_{s=1}^{\infty}\rho_i^s r_{i,t+1+s} \tag{1}$$

In the equation above, $r_{i,t+1}$ is the log return of stock i between time t and time $t+1$, $g_{i,t+1+s}$ is its log dividend growth rate ($d_{t+1+s} - d_{t+s}$ where d_t is the log dividend paid at time t) and the parameter ρ_i is related to the average price dividend ratio and is close to but smaller than one. While Campbell (1991) uses Equation 1 to study a market portfolio, it also holds for other portfolios and individual stocks since it follows directly from a dividend discount model.

Therefore, under rational expectations, three components drive a stock's realised return between time t and $t+1$: the expected return for this period at time t, and changes in dividend growth expectations, as well as changes in expected future return between time t and $t+1$. All else equal, an increase in expected future dividend growth will cause the realised return to be higher than expected, and an increase in future expected/required returns will cause it to be lower than expected. Importantly, under rational expectations, changes in expected dividend growth and expected future required returns can only be caused by new information arriving at time $t+1$ and cannot be predicted at time t. Hence, any potential return predictability must stem from differences in $\mathbb{E}_t(r_{i,t+1})$ (Bordalo et al., 2023).

Next, I transform the stock-level decomposition into a portfolio-level decomposition. Equation 1 can not only be applied to individual stocks but also to portfolios, by first taking the (potentially weighted) averages of the relevant variables across the stocks in the portfolios and then applying the log linearisation.

By applying Equation 1 separately to the aggregated long and short portfolios and taking the difference we can write:

$$r_{t+1}^{LS} \approx \mathbb{E}_t(r_{t+1}^{LS}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^{LS})^s g_{t+1+s}^{LS} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^{LS})^s r_{t+1+s}^{LS} \tag{2}$$

In the equation above, r_{t+1}^{LS} is the long-minus-short return of a predictor characteristic between time t and $t+1$ and g_{t+1}^{LS} the average log dividend growth rate between time t and $t+1$ in the long portfolio minus that of the short portfolio. Moreover, ρ^{LS} is a constant that is close to but smaller than one. Technically, a different value of ρ needs to be applied to the returns and dividend growth rates in the long portfolio than to those in the short portfolio, making it impossible to aggregate across portfolios. However, since ρ is always close to one by the construction of the Campbell-Shiller log linearisation, it is similar in both portfolios and we can approximate the ρ of each portfolio by ρ^{LS}. The resulting small approximation error is of no concern because I only use Equation 2 to qualitatively discuss the impact of the individual return components in order to give an intuition for my empirical specification.

In Equation 2, $\mathbb{E}_t\left(r_{t+1}^{LS}\right)$ captures the difference in expected/required returns between the long and the short portfolio, that is, the differences in compensation for risk. Under rational expectations, this is the only predictable component of Equation 2. Therefore, if a predictor is a risk factor and not associated with biased expectations, its ability to predict returns stems exclusively from its ability to predict $\mathbb{E}_t\left(r_{t+1}^{LS}\right)$. However, if we allow for biased expectations (denoted by $\tilde{\mathbb{E}}$), then both $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ and $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ are potentially predictable. If a characteristic can indeed predict either of these terms, this implies that its ability to predict returns is not (fully) driven by its association with risk ($\mathbb{E}_t\left(r_{t+1}^{LS}\right)$) but is at least partially explained by predictably biased beliefs. Since bias in the dividend growth expectations implies that $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s}) \neq [\mathbb{E}_{t+1} - \mathbb{E}_t](g_{t+1+s})$, realised returns will differ from their counterparts under rational expectations, indicating that the long-minus-short portfolio is mispriced. By a similar argument, bias in $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ also implies mispricing. This gives Proposition 1:

Proposition 1 *A return predictor characteristic is associated with mispricing if and only if it can predict either* $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ *or* $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ *for any* $s \in (0, \infty)$*, in addition to returns.*

Notably, the ability to predict changes in future expected dividend growth or required returns does not mean that a characteristic cannot also predict $\mathbb{E}_t(r^{LS}_{t+1})$ and is therefore associated with risk. However, since all predictors that are associated with biased expectations suggest profitable trading strategies and are evidence of market inefficiency, regardless of whether they are also associated with risk, I call all of these predictors mispricings.

Neither $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](g_{t+1+s})$ nor $[\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t](r_{t+1+s})$ are observable. However, for the former, forecasts by professional stock analysts can be used as a proxy. Therefore, I focus on the predictability of changes in dividend growth for the rest of the paper. Finding such predictability is a sufficient but not a necessary condition to identify that a predictor's returns are (partially) driven by mispricing because mispricing may also stem from the predictability of changes in future required returns or frictions which are not modelled in the Campbell-Shiller decomposition.

2.3 Dataset and Variable Construction

This section describes the data I use to classify return predictors into risk factors and mispricings. My primary dataset consists of 212 predictor portfolios from A. Y. Chen and Zimmermann (2021). I use their 2023.8 data release, which includes data up to the end of 2022. All included predictors showed evidence of return predictability in their original paper. I provide detailed descriptions of all predictors mentioned by name in this paper in Appendix E. The portfolio data includes monthly stock returns from CRSP as well as data on the predictor characteristic. I merge the portfolio data with earnings forecasts from the

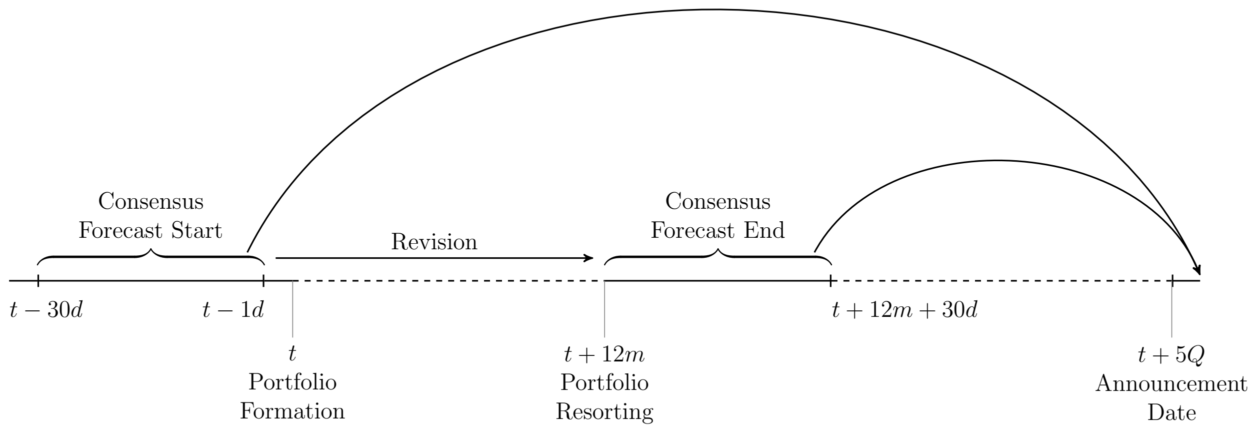
IBES unadjusted detailed history file, adjusted for stock splits using CRSP data. I use IBES data up to the end of 2022.

I use earnings forecasts instead of dividend forecasts due to their better data availability. These variables are closely related, as all earnings must eventually be paid out over a firm's complete life cycle. In fact, earnings forecasts may be a better proxy of the market's dividend growth expectations since I focus on forecasts for a single horizon for practical reasons. Arguably, a single earnings forecast is more informative about the firm's future dividend growth than a single dividend forecast, as earnings reflect the ability to pay dividends and are unaffected by pay-out policies such as dividend smoothing. Focussing on a single horizon is also sufficient to establish if a return predictor is associated with mispricing by Proposition 1. Finally, I use levels instead of growth forecasts. This can be done because growth rates are derived from levels, and under rational expectations, forecast revisions of any kind are unpredictable.

Figure 1 illustrates how I construct the forecast revision variable, which I will use as a proxy for the market's changes in dividend growth expectations. First, I calculate the median (consensus) five quarters ahead forecasts at the time of portfolio formation t using all forecasts made between 30 days $(t-30d)$ and one day $(t-1d)$ before portfolio formation. Then, I subtract this from the consensus forecasts for the same earnings announcement, including forecasts made within 30 days after the next portfolio resorting. Stocks are resorted into portfolios every one to twelve months in my sample of predictors. Hence, between zero and four announcements can happen between two sortings. Thus, the forecast made after resorting is for between one and five quarters ahead

The methodology above has the downside of only yielding one data point for each firm per resorting. Thus, I would only obtain one data point per year for predictors with yearly resorting. To increase my number of observations, I slightly adjust the procedure to generate data on a monthly frequency. Since predictors only have (confirmed) predictive power between the time of portfolio formation and the next resorting, I keep the date for the sec-

Figure 1: Timing of the forecast revisions



The figure illustrates the timing of the forecasts involved in the forecast revisions at the example of a predictor with yearly portfolio resorting. t is the date of portfolio formation. d stands for days, m stands for months and Q stands for quarters.

ond forecast fixed at the time of resorting. The date of the first forecast is changed on a monthly frequency. Since the return predictor is linked to abnormal returns across the entire period between portfolio formation and resorting, a mispricing predictor should also predict revisions between any two forecasts made in this period.

Both the initial forecast and the one at the time of the next resorting are for the same target date. The initial forecast is always for five quarters ahead. Since the date at which the initial forecast is made moves every month, whereas the date of the second forecast only changes when a resorting occurs, this methodology implies that the difference in forecast horizon between the initial forecast and the forecast at the end of the resorting period changes over time.

To illustrate this, Table 1 shows an example of a predictor with yearly resorting on the first of June and for a firm with earnings announcements on the first of January, April, July and October, respectively.

In Appendix D, I show that my results are broadly similar but less significant if I do not use this procedure. Moreover, I compare my results for predictors that have a monthly

Table 1: Overview of Forecast Matching

Date initial forecast made	Horizon initial forecast	Target date (both forecasts)	Date forecast after resorting	Horizon forecast after resorting
1st June year t	5Q	1st July Year t+1	1st June year t+1	1Q
1st July year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st August year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st September year t	5Q	1st October Year t+1	1st June year t+1	2Q
1st October year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st November year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st December year t	5Q	1st January Year t+2	1st June year t+1	3Q
1st January year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st February year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st March year t+1	5Q	1st April Year t+2	1st June year t+1	4Q
1st April year t+1	5Q	1st July Year t+2	1st June year t+1	5Q
1st May year t+1	5Q	1st July Year t+2	1st June year t+1	5Q

The table illustrates the matching of forecasts used to compare forecast revisions in the predictor long and short portfolios at the example of a predictor with yearly resorting on the first of June and for a firm with earnings announcements happening on the first of January, April, July and October, respectively.

resorting frequency and, therefore, do not require this procedure and those with yearly resorting. I find that the results are very similar, alleviating concerns that the procedure may introduce a bias in my results.

I choose five quarters ahead forecasts since this is the shortest horizon for which I can have a forecast for the same announcement date at portfolio formation and after the next resorting for resorting periods of 12 months, which is the longest period I study[7]. The choice of a short horizon has the upside of better data availability and the downside that long-term growth expectations (which are also available from IBES) are theoretically the more important driver of stock returns. I may, therefore, fail to capture some mispricing predictors by focusing on less important shorter horizons. In Appendix C, I repeat my main analysis with long-term growth (LTG) expectation data.

I calculate the forecast revision of the consensus forecast for stock i at month t with x month until the next resorting for the earnings announced at date d as:

$$forecast\,revision_{i,d,t,x} = \frac{forecast_{i,d,t+x} - forecast_{i,d,t}}{price_{i,t-2y}} \tag{3}$$

I follow the standard practice and scale forecast revisions (and any other EPS-related variables) by price[8]. I use a 2-year lagged price because otherwise mispricing can affect my

[7]There is a single predictor with a 36-month resorting period in my sample which I exclude.

[8]The reason for this scaling is that a company's earnings per share can only be meaningfully interpreted

results through scaling. If a stock is overpriced, its forecast revision will be scaled by a larger price and hence be smaller than the revision of an otherwise identical stock that is correctly priced or underpriced. Since I study mispricing-driven return predictability, there will be differences in the level of mispricing between the stocks with predictably higher returns and those with predictably lower returns at time t by design. However, it is unlikely that these differences extend two years back because return predictability generally does not last that long.

In addition to scaling by price, I winsorize all variables derived from EPS forecasts. Winsorization is done based on the interquartile range (IQR) method. Let $Q1$ be the value of the first quartile and $Q3$ the value of the third quartile. I winsorize all data below $Q1 - 1.5(Q3 - Q1)$ or above $Q3 + 1.5(Q3 - Q1)$. I do this because even after scaling by the price, some firms still have significantly larger EPS, which can cause them to dominate my analysis quantitatively.

In addition to forecast revisions, parts of my analysis also use forecast errors defined as:

$$forecast\,error_{i,d,t} = \frac{actual\,value_d - forecast_{i,d,t}}{price_{i,t-2y}} \tag{4}$$

2.4 Multiple Hypothesis Testing

Since I am testing significance for a large set of predictors, multiple hypothesis testing is a concern. There are two common approaches to address this issue: controlling the false discovery rate (FDR) and controlling the family-wise error rate (FWER). Controlling the FDR limits the share of false positive results to α (i.e., 5% for the most common value of α). Controlling the FWER limits the probability of a single false positive among all hypothesis α. Controlling the FWER is stricter than controlling the FDR, particularly for a large number

in conjunction with its share price because EPS crucially depend on the number of shares outstanding and can easily be changed by (reverse) stock splits. This can cause two otherwise identical companies to have vastly different earnings per share if they have different amounts of outstanding shares. Scaling EPS-related variables by price transforms them to earnings per $ of market capitalisation, which is unaffected by the number of shares outstanding and is only driven by economically meaningful measures.

of tests, since for the FDR, the number of acceptable false positives scales with the number of tests, whereas controlling the FWER tries to rule out any false positives regardless of the number of tests. To achieve this, controlling the FWER requires higher and higher critical t-values as the number of tests increases, leading to a large fraction of false negatives.

In my main analysis, I report results without MHT controls and results that control the FDR using the method by Benjamini and Hochberg (1995). I chose to control the FDR and not the FWER because I am trying to provide an accurate picture of the fraction of mispricings related to biased expectations among all predictors. Therefore, I am not only concerned about classifying a predictor as mispricing by mistake (type one error) but also about failing to classify a predictor as mispricing (type two error). Since controlling the FDR strikes a balance between limiting both error types while controlling the FWER tries to rule out type one errors at the expense of allowing more type two errors, controlling the FDR is more suited to my research question.

2.5 Descriptive Statistics

Table 2 shows descriptive statistics for my main variables, aggregated across the 212 predictors, that showed significant evidence of return predictability in the original sample. To calculate the aggregated descriptive statistics, I first calculate each variable's mean and standard deviation separately for each predictor and then take the mean of those means and standard deviations across the 212 predictors. Only the values for N, the number of data points with available forecast revisions per predictor, are calculated differently. This variable is already defined on the predictor level. Hence, the mean and standard deviation across predictors can be calculated directly. As expected, the return is substantially larger in the long portfolios (5.01%) compared to the short portfolios (4.10%). Forecasts are also more positive in the long portfolio, both at portfolio formation and after the next resorting. Forecast revisions are negative in both portfolios but more so in the short portfolio, suggesting that there are at least some mispricing predictors in my sample. As discussed

Table 2: Descriptive Statistics

	long			short		
	mean	median	sd	mean	median	sd
Return	5.01%	2.19%	31.42%	4.10%	1.65%	30.20%
Forecast Start	0.017	0.016	0.018	0.015	0.013	0.017
Forecast End	0.016	0.015	0.018	0.014	0.012	0.017
Forecast Revision	-0.00076	-0.00026	0.00525	-0.00091	-0.00033	0.00509
Forecast Error Start	-0.00328	-0.00158	0.01222	-0.00362	-0.00180	0.01187
Forecast Error End	-0.00213	-0.00077	0.00975	-0.00234	-0.00091	0.00944
Share Forecast Available	13.94%	13.97%	5.93%	11.93%	12.05%	5.39%
N	52969	41738	65314	56483	44668	56187

This table shows descriptive statistics for the main variables using return data from 1984 to 2022 and earnings (forecast) data from 1984 to 2022. It includes all data points with available forecasts. Forecast-related variables are scaled by price and winsorized. The data is separated into long and short portfolios. For all variables except N, I first calculate the mean and the standard deviation of the variable for each portfolio of each predictor and then aggregate across predictors by taking the mean of the individual means and standard deviations. The variable N is already defined on the predictor level, so the mean and standard deviation across predictors can be calculated directly. *Return* is the monthly stock return. *Forecast Start* and *Forecast End* are the consensus forecast values at portfolio formation and after the next resorting, respectively. *Forecast revision* is defined as *Forecast end* minus *Forecast Start*. *Forecast Error Start* and *Forecast Error End* are the actually announced values minus the consensus forecast values at portfolio formation and after the next resorting, respectively. *% Forecast Available* is the number of return observations with available forecast revisions divided by the total number of return observations. A more detailed description of the variable construction can be found in Section 2.3.

above, forecast data availability is an issue. On average, I have forecast revisions for slightly more than 10% of the data points for which I have return observations. This will reduce the statistical power of the forecast analysis relative to the return analysis and can lead to some mispricings not being detected.

3 Mispricings

3.1 How Many Predictors are Mispricing?

As a first step, I test which predictors offer a significantly positive long-minus-short return using the data up to the end of 2022 (i.e., including the post-publication period). To do this, I calculate the monthly portfolio return using equal or value weighting depending on which the original paper used. Next, I do t-tests to determine if the long minus short return is significantly larger than zero with a t-statistic above 1.96 and find that this is the case for 172 predictors. Since significant return predictability is a prerequisite for a predictor, I only use these predictors in the subsequent analysis. I do not control for multiple hypothesis testing in this first step because doing so would cause me to lose predictors even if their level of return predictability is the same as in the original paper since the original papers generally do not use MHT controls. If this approach causes a few false predictors to pass the initial test, this will likely cause me to underestimate the share of mispricings among the true predictors in my later analysis, as spurious predictors are unlikely to be classified as mispricing by my test. In Table A2 in Appendix D, I show that my results are similar using only predictors that have significantly positive long minus short returns after controlling the false discovery rate.

By Proposition 1, a predictor is unable to predict revisions in dividend (or earnings) expectations if it is a risk factor but may do so if it is a mispricing because predictable changes in earnings expectations are one of the two channels through which biased beliefs may drive return predictability. I, therefore, say that a predictor represents mispricing if I find evidence that it predicts forecast revisions, which I test by running the following regression using monthly-level data:

$$forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x} \tag{5}$$

In the regression, i indicates a firm, d indicates an announcement date, t is a month, and

x is the number of months from t to the following portfolio resorting. $long_{i,t}$ is a dummy that is one if stock i is sorted into the long portfolio at time t and zero if it is sorted into the short portfolio. If a stock is neither in the short nor the long portfolio, it is excluded from the regression. Finally, $X_{i,t}$ is a vector of potential control variables.

A positive and significant coefficient of $long_{i,t}$ indicates that the predictor reflects mispricing since, for risk factors, there should be no difference in expectations between the long and the short portfolio. If the coefficient of $long_{i,t}$ is not positively significant, this suggests that the predictor is a risk factor or a mispricing related to biased beliefs about future required returns. It could also mean that the analysis lacks statistical power to uncover changes in expectations in the noisy analyst forecast data or that earnings expectations, for a horizon I do not study, drive the predictor's abnormal returns. A crucial assumption behind this regression is that forecasts by stock market analysts are a valid proxy of the expectations of market participants. Previous research has found that data on analyst expectations can be used to explain patterns in stock returns and asset prices (see Adam and Nagel (2023) for a review), suggesting that they can indeed be used to proxy for the market's expectations.

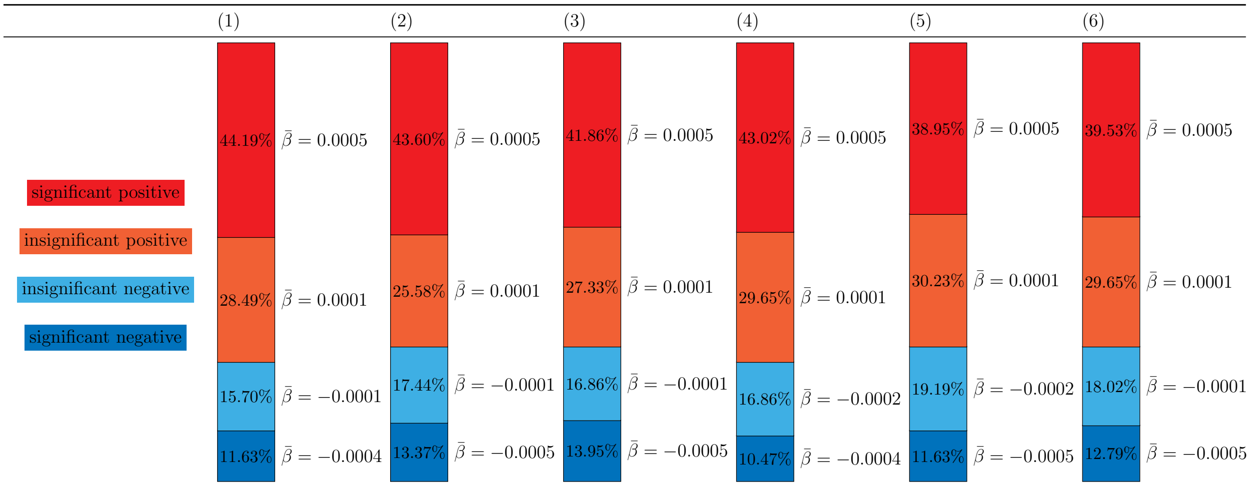
Another potential concern is that analysts (partially) extrapolate from returns when revising their forecasts (as suggested by, , Ben-Rephael et al. (2021)). This could cause forecasts to be different in the long and the short portfolio, even for risk factors. I can address this issue by controlling for the return during the time between forecast revisions. However, since I form a consensus forecast by aggregating individual forecasts made at different points in time, I cannot perfectly control for the return information that was available to each individual analyst when they made their forecast, which means that there is some concern left that by results are biased upwards by analysts who extrapolate from returns. In Section 3.3, I provide evidence that return extrapolation by analysts cannot explain my results.

Controlling for returns also introduces a downward bias in my result as it filters out firm-relevant news incorporated into both forecasts and prices. Nevertheless, existing evidence suggests that analysts add private information and do not just extrapolate returns van

Binsbergen et al. (2022), which suggests that substantial variation will be left in the analyst forecast even when controlling for returns.

I run separate regression for each predictor with significantly positive long minus short returns in my data. The aggregated results are shown in Table 3.

Table 3: Share of Mispricings among the Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	50568	50568	50568	50568	50568	50568
Number of Regressions	172	172	172	172	172	172

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

The share of predictors that have a significant positive long dummy and are classified as mispricing varies from around 39% to around 44%. It is highest in the baseline specification (1), with no controls or fixed effects and no adjustment for multiple hypothesis testing. Interestingly, controlling for the return does not matter much: between specifications (1) and (2), the share of significantly positive long dummies only drops by less than one percentage point. This suggests that analysts learning from prices is not a significant concern. Column (3) adds year/month fixed effects relative to column (2), which is useful for two reasons: First, due to limited data availability, the number of firms in the long and short portfolio varies between months, which could bias my result upwards if I have more (less) stocks in a month with generally good (bad) news in the long portfolio relative to the short portfolio. Second, due to the way the forecast revision is constructed (described in Section 2.3), the difference in forecast horizons between the initial forecast and the revised forecast varies between months and the year/month will absorb any potential effect of this difference. Including fixed effects decreases the share of significant strategies by less than two percentage points. I do not include firm fixed effects, which is commonly done in regressions with panel data because, for many return predictors, the differences between firms are the reason for return predictability and not something I want to control for. For instance, the size (market capitalisation) of the firm is one common return predictor, and including firm fixed effects would almost fully absorb the effect of the firm size since firms that are in the smallest decile in some periods generally do not move to the largest decile in other periods. Many other accounting characteristics are similarly persistent.

Columns 4,5 & 6 are identical to columns 1,2 & 3, respectively, except that they control for multiple hypothesis testing (MHT) by keeping the false discovery rate at 5%. Naturally, this leads to a lower share of significant strategies. However, the effect is less than five percentage points for all specifications, suggesting that the significant long dummies generally have t statistics comfortably above the standard threshold of 1.96.

The coefficients are remarkably stable across specifications. Among the strategies with a significant positive long dummy, the average coefficient is always 0.0005. The most straightforward way to interpret the magnitude of the coefficient in the context of the research question of this paper is to test how much of a return difference between the long and the short portfolio is implied by the higher earnings forecast revisions in the long portfolio exemplified by the positive coefficient. By comparing the implied return difference to the actual return difference, I could assess if biased expectations drive the entire return predictability of mispricing predictors.

Unfortunately, to do so accurately, I would need to study changes in earnings expectations for all future earnings announcements because, by Equation 1, I need a proxy for all future dividend growths. Since I have to focus on a single horizon for practical reasons, I can only do a rough back-of-the-envelope calculation in Appendix B. The result suggests that the more positive changes in earnings expectations in the long portfolio are sufficient to explain the full return differential between long and short portfolio stocks and, if anything, suggest an even larger return differential.

Interestingly, around 13% of predictors also have significant negative long dummies, which is predicted for neither mispricings nor risk factors. Since my controls for multiple hypothesis testing limit the rate of false positives to 5%, this is unlikely to be purely a statistical artefact. While this result is not predicted for mispricings, it is also no contradiction. Since returns are driven by changes in expectations for all future time periods, it is possible that expectations for a single period change in the opposite direction. In contrast, a risk factor should not be able to predict forecast revisions for any time period and in any direction. Hence, one potential interpretation for the result is that the associated predictors predict lower short-term but higher long-term earnings. Another potential explanation is that the predictor is associated with both risk and mispricing but in the opposite direction. If the predictor's long portfolios are riskier but also overpriced at portfolio formation and the compensation for risk quantitatively dominates the mispricing effect, returns would be higher in the long portfolio relative to the short portfolio while expectation revisions in the long portfolio are more negative than those in the short portfolio at the same time. In this case, investing in the short legs of the predictor portfolios would offer superior risk-adjusted returns despite the lower raw return of these portfolios.

In Appendix C, I repeat my analysis using long-term growth forecasts, which capture the analysts' growth expectations over a firm's full business cycle. I find that 18% of predictors have significantly more positive LTG forecast revisions in the long portfolio compared to the short portfolio. This is around half as many predictors as with quarterly data, which is likely at least partially driven by lower statistical significance because of lower data availability. Overall, 22 predictors have significantly more positive forecast revisions in the long portfolio both with LTG data and with quarterly data and for nine predictors the difference is only significant with LTG data. Importantly, predictable expectations revisions at any horizon imply mispricing, so these results suggest that my classification based on just quarterly data slightly underestimates the share of mispricing among the return predictors.

My sample contains predictors with different resorting periods. For roughly half of the predictors, the stocks are resorted into portfolios every month. The other predictors are almost all resorted yearly, while a few predictors also have 3-monthly and six-monthly resorting periods. In Figure A1 in Appendix D, I examine if the share of detected mispricings among the predictors varies with the resorting frequency and find that it is broadly similar for predictors with monthly and yearly resorting. Surprisingly, all predictors with three-monthly and six-monthly resorting are mispricings, but there are not enough data points to draw a firm conclusion from this result.

The return predictability of each predictor in my dataset was tested separately. Many predictors are similar to each other, suggesting that some predictors are likely redundant. Feng et al. (2020) address the issue of which predictors to select from a large set of candidates using a stepwise procedure. They start with the four predictors from the Carhart (1997) four-factor model: Market Beta, Size, Book-to-Market and Momentum. Then, they test for each predictor in their sample of 150 predictors how much it adds to the four-factor model. In the next step, they add the predictor that added most to the existing predictors to the model and test all remaining predictors against the new model with five factors. They repeat this procedure until no more predictor adds significantly to the model. Using this procedure, they select 26 predictors, including the original four that were used as a starting point. Out of those, I have 18 in my data, 14 of which offer significant long minus short returns.

For each of these 14 predictors, I examine if it is linked to mispricing. Table 4 shows that four have a t-statistic above 1.96 and are, thus, classified as mispricing by my method.

Table 4: Mispricing test for the stepwise selected predictors of Feng et al. (2020)

name	T-value	Parameter	Standard Error	Description
Accruals	1.25	0.0002	0.000139	Accruals
BMdec	-0.34	-0.0001	0.000162	Book to market using December ME
ChInvIA	0.40	0.0000	0.000101	Change in capital inv (ind adj)
DelFINL	3.29	0.0003	0.000087	Change in financial liabilities
Illiquidity	-3.64	-0.0008	0.000218	Amihud's illiquidity
Mom12m	7.15	0.0007	0.000103	Momentum (12 month)
NOA	2.21	0.0001	0.000050	Net Operating Assets
Size	-7.87	-0.0009	0.000112	Size
Tax	5.81	0.0006	0.000108	Taxable income to income
VolSD	-0.60	-0.0000	0.000079	Volume Variance
AdExp	1.77	0.0003	0.000168	Advertising Expense
SP	-0.44	-0.0001	0.000183	Sales-to-price

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor that survived the stepwise selection procedure from Feng et al. (2020), is included in my dataset and has significant long minus short returns in my data. The regression controls for the return over the period between the two forecasts included in the revision and uses Year/Month fixed effects. The reported t-values, parameters and standard errors refer to $long_{i,t}$. Standard errors are clustered on the firm and the year/month level.

3.2 Which Predictors are Mispricings?

In this section, I take a closer look at which types of predictors are related to mispricing. First, I examine if the magnitude of the long-short return of the predictors differs between mispricings and other predictors. The mean return of the mispricing predictors is 40 basis points and that of the other predictors is 38 basis points. Thus, the returns are remarkably similar. Table 5 shows the mispricing classification of the 20 predictors with the largest long-minus-short returns. We can see that out of the seven predictors with the largest long-minus-short returns, only 1 is classified as mispricing, but among the top twenty predictors, thirteen predictors are classified as mispricing. This suggests that mispricing is common among predictors with large long-minus short returns, except at the very top.

Table 5: Mispricing Classification of the Predictors with the Largest Long-Short Returns

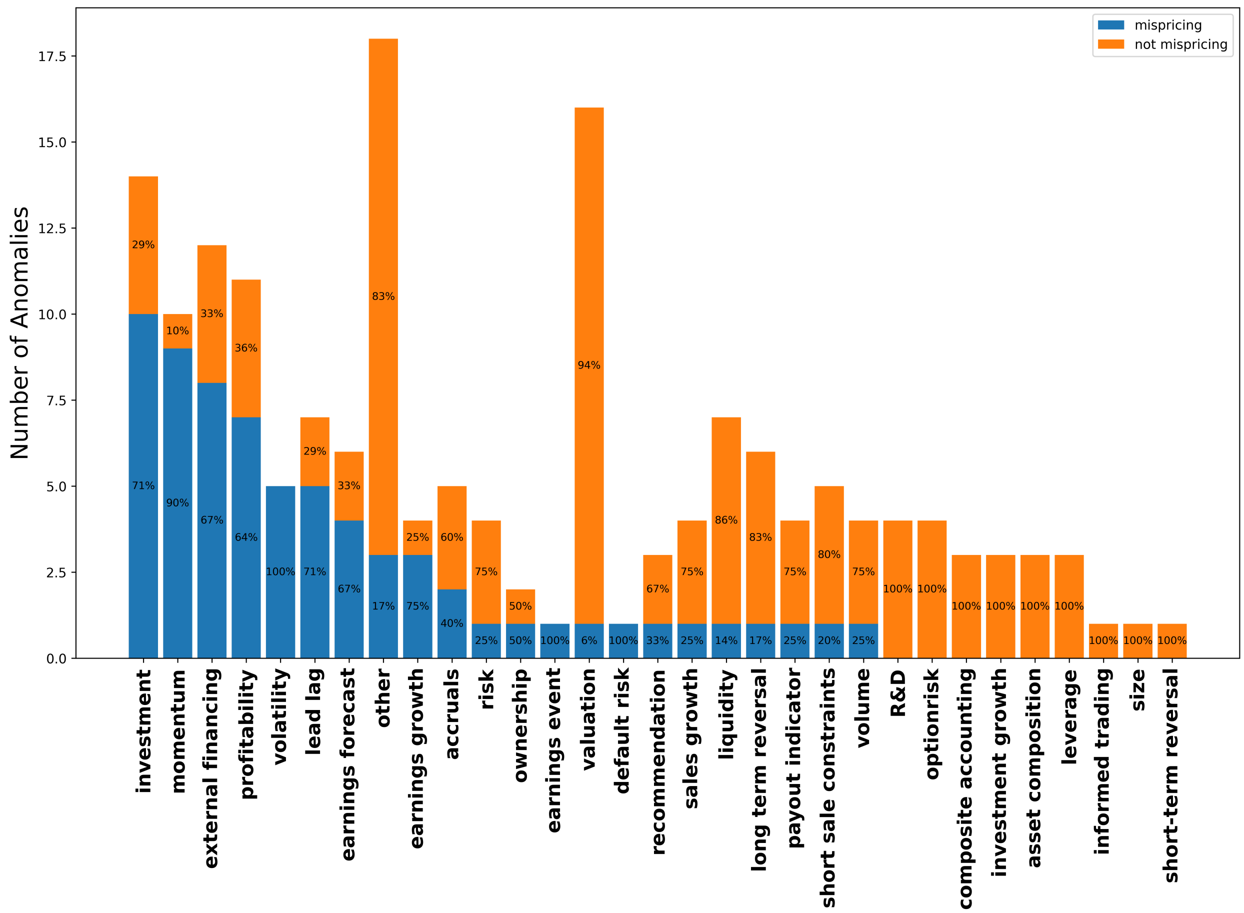
Name	Description	LS Return	Mispricing
IO_ShortInterest	Inst own among high short interest	3.20	No
STreversal	Short term reversal	2.74	No
TrendFactor	Trend Factor	1.66	No
IndRetBig	Industry return of big firms	1.36	Yes
Frontier	Efficient frontier index	1.32	No
AccrualsBM	Book-to-market and accruals	1.32	No
SmileSlope	Put volatility minus call volatility	1.27	No
FirmAgeMom	Firm Age - Momentum	1.26	Yes
roaq	Return on assets (qtrly)	1.19	Yes
IntMom	Intermediate Momentum	1.14	Yes
XFIN	Net external financing	1.10	Yes
AnnouncementReturn	Earnings announcement return	1.09	Yes
MomVol	Momentum in high volume stocks	1.09	Yes
retConglomerate	Conglomerate return	1.08	Yes
MomOffSeason	Off season long-term reversal	1.03	No
Mom6mJunk	Junk Stock Momentum	0.97	Yes
ChTax	Change in Taxes	0.95	Yes
EarningsStreak	Earnings surprise streak	0.94	Yes
AssetGrowth	Asset growth	0.94	Yes
Mom12mOffSeason	Momentum without the seasonal part	0.94	Yes

The table shows the long minus short return of the 20 return predictors with the largest long minus short return. It also shows the result of a mispricing classification based on the regression of forecast revision on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Column 6 in Table 3) for these predictors.

Next, I group the predictors into economic categories based on (A. Y. Chen and Zimmermann, 2021) and examine which types of predictors are most likely to be classified as mispricing. I also compare these interpretations of the different predictor types in the literature.

Figure 2 shows the share of mispricing predictors in each category using the regression specification that controls for return, includes fixed effects and controls the false discovery

Figure 2: Share of Mispricings by Predictor Category



The figure shows the share of predictors that are classified as mispricing by predictor category. The categories are taken from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification as mispricing is based on the regression of the forecast revision on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Column 6 in Table 3).

rate (Column 6 in Table 3). An individual-level overview of all mispricing predictors can be found in Table A7 in Appendix D. Reassuringly, a large fraction of *lead-lag* predictors, which are supposed to capture delayed reactions to news is classified as mispricing.

Momentum is the category with the highest share of mispricing predictors (nine out of ten), suggesting that momentum is caused by under or overreaction to news and not by risk. This is generally in line with the interpretation of the original authors who discovered the return predictors, who also interpreted their result as mispricing or did not provide an interpretation. Moreover, *investment*, *profitability* and *external financing* predictors are mostly classified as mispricing. The majority of these predictors were also classified as mispricing by the original authors, but there are also several predictors in each category that were originally interpreted as risk.

Importantly, the version of momentum predictor used in the Carhart (1997) four-factor model and the versions of the profitability and investment predictors used in the five-factor model by Fama and French (2015) are linked to mispricing[9].

Finally, all *volatility* predictors are classified as mispricing. Three of these come from Ang et al. (2006), who interpret their results as risk. The other volatility measures are the earnings forecast dispersion from Diether et al. (2002) and the maximum return over the last month from Bali et al. (2011) who both interpret their results as mispricing.

There are eleven categories with more than one significant predictor for which only one predictor is classified as mispricing and five where no predictor is classified as mispricing. Valuation stands out as a category with 16 significant return predictors only one of which is classified as mispricing. This is in line with the results of Han et al. (2020), who combine several valuation predictors into a single score and find this combined predictor is not linked to mispricing. Overall, these results suggest that the mispricing predictors are concentrated in certain categories of predictors.

3.3 Detecting Spurious Predictors

So far, I have assumed that all return predictors in my dataset are either risk factors or mispricings. However, It is also possible that some are not valid return predictors at all but instead the result of data mining or multiple hypothesis testing. This is a potential concern for my methodology. Recall that by Equation 1, a stock has a high return between t and $t+1$

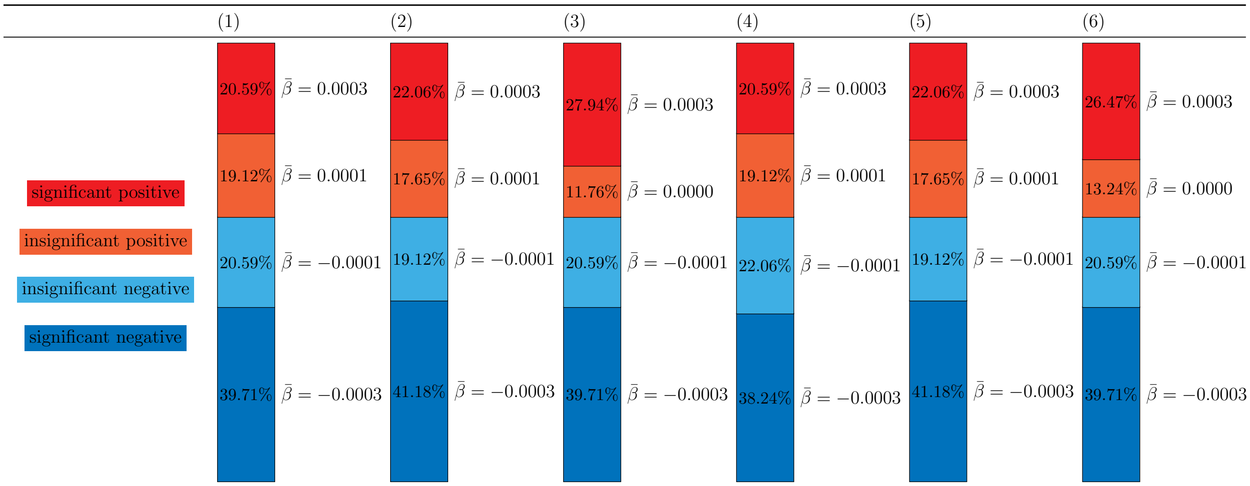
[9] The momentum predictor is the return in the preceding 12 months following Jegadeesh and Titman (1993). The Investment predictor is defined as the change in total assets between year t-2 and year t-1, which is originally from Cooper et al. (2008) and the profitability factor is defined as revenue minus cost administrative expenses interest expenses, scaled by book value of equity and follows Fama and French (2006).

for three possible reasons: either its required return for this period is high, or the market revised its expectation for the discounted sum of future dividend growths upwards or because the market revised its future required returns downwards. If stocks in the long portfolio of a spurious predictor have higher in-sample returns than those in the short portfolio, this will likely be driven by a mixture of these three reasons. Therefore, these stocks likely have more positive dividend growth expectation revisions on average. Importantly, these more positive revisions do not imply mispricing, as they are not predictable. Instead, stocks with good (bad) firm-related news are selected ex-post to be in the long (short) portfolio, as also discussed by (Engelberg et al., 2018).

Asymptotically, if returns and revisions of earnings forecasts are correlated ex-post for at least some stocks in a spurious predictor's long or short portfolio, this will cause my empirical specification to misclassify such a spurious predictor as mispricing. However, in practice, two factors make such a misclassification less likely. First, I control for the return in my main specification, which filters out a mechanical correlation between returns and analyst forecast revisions. Second, changes in dividend expectations are only one reason why a stock could have higher returns in a given period. Therefore, the long (short) portfolio of a spurious predictor will include only some stocks that have biased dividend expectations, whereas other stocks have high (low) risk or are mispriced for other reasons. In contrast, if a valid mispricing predictor is driven by biased expectations, the majority of stocks in its long portfolio should see more positive dividend expectation revisions than the stocks in its short portfolio. With limited statistical power, it is thus much less likely that a spurious predictor is classified as mispricing.

I do two robustness tests to study how likely it is in practice that a spurious predictor gets classified as mispricing. First, I apply my test to a set of 66 predictors from A. Y. Chen and Zimmermann (2021) that had no significant evidence of return predictability both in the original paper that tested them and in the extended dataset using data up to the end of 2022. In this sample, around 26% of predictors are classified as mispricing. While this is less than in the sample of significant return predictors, it highlights that some potentially spurious predictors are classified as mispricing. However, each of these predictors was initially tested because the researchers had the hypothesis that it might predict returns, so there are likely some true return predictors that just do not pass the statistical significance hurdle in this sample.

Table 6: Detected Mispricings among Insignificant Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	62346	62346	62346	62346	62346	62346
Number of Regressions	68	68	68	68	68	68

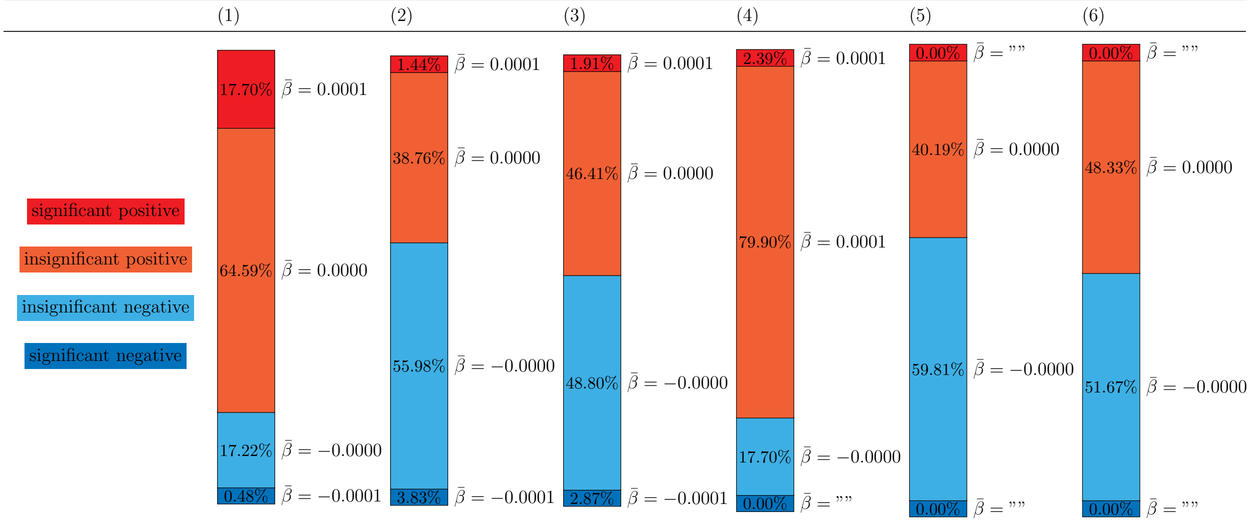
The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in more detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for 66 characteristics with insignificant evidence of return predictability, both in the original sample and in the extended sample, using all data up to the end of 2022. The table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

To assess how my test performs with a sample of predictors that are unambiguously spurious, I apply my methodology to a set of simulated spurious predictors. I generate these predictors by randomly sorting stocks into portfolios each month. I exclude the stocks with the worst 2% returns from the long portfolio and those with the best 2% returns from the short portfolio to ensure I can find portfolios with sizeable long-minus-short returns in a reasonable time. I retain a simulated predictor if its long minus-short return is above the 90th percentile of long-minus-short returns for actual predictors.

To ensure that my simulated predictors are representative of the real predictors, they are designed to match them on the resorting frequency, the share of stocks assigned to the long/short portfolio and whether returns are equal or value-weighted. For example, after removing some predictors with more complex assignment rules, which I cannot represent in my simulated data, 12% of the predictor portfolios in my main analysis have monthly resorting, are equal-weighted and assign the most extreme deciles into the long and short portfolio respectively. Consequently, also 12% of my simulated portfolios have these properties. I generate a number of simulated portfolios that is as close as possible to the number of real predictor portfolios (172) while maintaining the correct share of different types of predictors, giving me a target number of 209 portfolios.

I then run the regression specified in Equation 5 for each simulated portfolio. Table 7 shows the results, and its columns are directly comparable to those of Table 3, which uses the actual data. In the baseline specification, 17.70% of the simulated predictors are classified as mispricing. However, adding return controls reduces this share to 1.44% and even without return controls, only 2.39% of simulated predictors are classified as mispricing once multiple hypothesis testing is controlled for. In my main specification, with fixed effects, return controls and MHT controls, no simulated predictor is classified as mispricing, suggesting that my methodology is extremely unlikely to misclassify spurious predictors as mispricing. Moreover, the test also suggests that my results are not driven by analysts learning from returns. If analysts' forecast revisions were to mechanically follow returns, we would expect to see significantly more positive forecast revisions in the simulated long portfolios compared to the short portfolios since they have higher returns.

Table 7: Detected Mispricings among Simulated Spurious Predictors



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	54633	54633	54633	54633	54633	54633
Number of Regressions	209	209	209	209	209	209

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in more detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for 209 simulated predictor portfolios that are designed to be spurious, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

4 Build-up or Resolution?

4.1 How many Predictors are Build-up and Resolution?

In a recent paper, van Binsbergen et al. (2023) raise the question of whether mispricing predictors capture the build-up or the resolution of mispricing. They calculate the fundamental value of each firm in a predictor portfolio by discounting realised future dividends and capital gains with an estimated SDF. They then compare fundamental values to market prices to calculate price wedges, i.e., deviations of the stock price from the fundamental value. By examining how the price wedges line up with the abnormal returns of the predictor portfolios they can determine if a predictor is build-up or resolution. They find that around a third of predictors are build-up. In Section 3, I identified 72 out of my total 172 predictors that capture mispricing driven by biased earnings expectations (based on specification 3 of Table 3). In this section, I classify these predictors into build-up and resolution predictors using an approach that is based on forecast errors which is complementary to the price-wedge-based approach.

For build-up predictors, prices are either correct at portfolio formation or the long portfolio is already overpriced relative to the short portfolio. The higher returns of the long portfolio lead to a (further) divergence of prices from fundamental value over the return period. This divergence can capture excessive optimism about stocks in the long portfolio, excessive pessimism about stocks in the short portfolio or both. In all of these cases, stocks in the long portfolio should be overpriced relative to those in the short portfolio after the period of abnormal returns. This should also be reflected in more optimistic earnings expectations for stocks in the long portfolio relative to stocks in the short portfolio at the end of the return period. By similar reasoning, for resolution predictors, the long portfolio is underpriced relative to the short portfolio at the time of portfolio formation. This implies that earnings expectations are more negative for stocks in the long portfolio than stocks in the short portfolio.

Finding evidence of biased expectations at the time of portfolio formation (at the end of the return period) is insufficient to classify a variable as build-up (resolution). Another requirement is that changes in expectation during the return period have led to a divergence (convergence) of expectations. However, since I only classify predictors as mispricing if they had more positive forecast revisions for stocks in the long portfolio than for stocks in the short portfolio, this is the case for all mispricings. Consequently, it is sufficient to focus on the forecast errors in this section.

To determine if a variable reflects the resolution of mispricing I, therefore, run the following regression:

$$forecast\,error_{i,d,t} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t} \tag{6}$$

To determine if a variable reflects the build-up of mispricing, I run the following regression:

$$forecast\,error_{i,d,t+12m} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t} \tag{7}$$

In the regression formulas, i indicates a firm, d indicates an announcement date, t is the month of a portfolio formation and m stands for a month. $long_{i,t}$ is a dummy that is one if stock i is sorted into the long portfolio at time t and zero if it is sorted into the short portfolio. Unlike in the analysis in Section 3, which used monthly level data, I only use one data point per portfolio formation in this analysis. This implies that the data frequency depends on the resorting period and varies between monthly and yearly depending on the predictor. I do not use monthly data because, for resolution predictors, the difference in expectations between the long and the short portfolio diminishes over the return period. Hence, the further a monthly data point is away from the portfolio formation, the smaller the difference in expectations should be, and at the time of the next rebalance, the mispricing may be fully resolved. By a similar logic, there is initially no difference between expectations for the long and the short portfolio for build-up predictors, and the difference has only fully materialised at the end of the period for which the predictor predicts abnormal returns. Thus, I only have precise predictions for differences in forecast errors between the long and the short portfolios at the time of portfolio formation and after the next resorting. Hence, I cannot use monthly-level data.

As with the mispricing analysis, a potential issue with my methodology is that analysts may learn from returns. This issue is less severe here because I only use predictors classified as mispricing. Hence, provided that this classification is valid, there should be no predictors where the link between the predictor and the analysts' forecasts is purely driven by return extrapolation in this analysis. Nevertheless, extrapolation from returns would still affect the results. Therefore, I control for the return. For resolution predictors, I ideally want to control for the return over the horizon during which the mispricing that is being resolved has built up. Since this is unknown, I choose to control for the monthly returns of the 12 months preceding the forecast. For build-up predictors, mispricing is associated with the predictors' abnormal returns. Consequently, I control for the return between the formation of the predictor portfolio and the time the forecast was made.

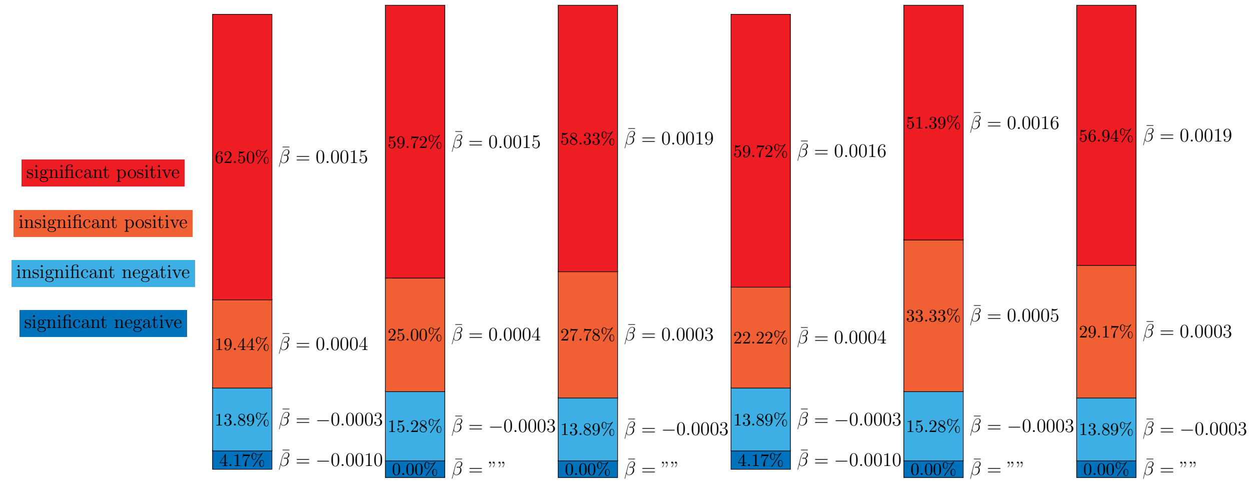
I use the predictors that were classified as mispricing using the specification with return controls and year/month fixed effects but without MHT control (Column 3 of Table 3). I do not control for multiple hypothesis testing for the same reason I do not do it when assessing the return predictability of the predictors to narrow down my dataset for the mispricing analysis: false negatives and false positives both negatively affect the quality of my dataset. Hence, it does not make sense to focus excessively on controlling the rate of false positives.

Table 8 shows the results of a regression of forecast error on the day of portfolio formation on the long dummy. A positive long dummy indicates that analysts are more pessimistic about stocks in the long portfolio, which is predicted for resolution predictors. I find that, in my preferred specification (6), in which I control for the return, include year/month fixed effects and adjust for multiple hypothesis testing, 56.94% of mispricings are classified as resolution. All other specifications have broadly similar results.

The share of strategies with a significant positive coefficient on the long dummy is low at between 4.17% and 0%, suggesting that there are few predictors where the long portfolio is already underpriced relative to the short portfolio at the time of portfolio formation.

To detect predictors that build up mispricing, I would ideally examine forecast errors at the end of the longest possible period for which return predictors can predict abnormal returns because this is the point where the mispricing is fully build up. Unfortunately, researchers who discover return predictors generally do not attempt to discover this period

Table 8: Share of Resolution Predictors

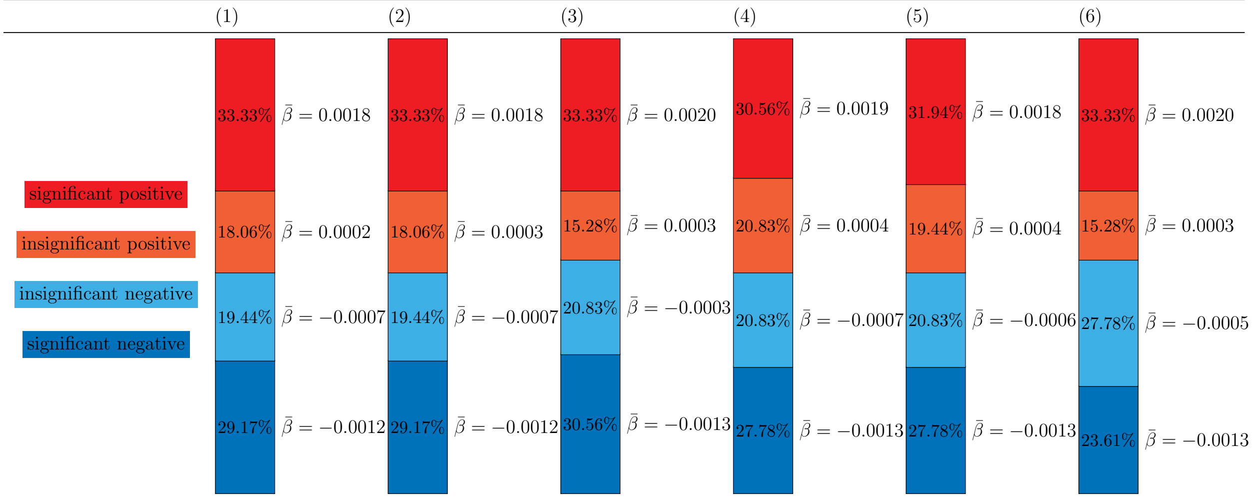


	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	61727	56040	56040	61727	56040	56040
Number of Regressions	72	72	72	72	72	72

The table shows the result of the regression $forecast\ error_{i,d,t} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t}$, where $forecast\ error_{i,d,t}$ indicates the consensus forecast error on the day of portfolio formation. The regression is run separately for all predictors identified as mispricing in the regression shown in Column 3 of Table 3 and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the 12 monthly returns before the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

and instead either test return predictability with monthly resorting of stocks into new portfolios or with yearly resorting. Therefore, I need to make an assumption about how long a given portfolio sort can predict abnormal returns. I choose to examine errors of five quarters ahead forecasts made one year after the portfolio formation because one year is the longest portfolio resorting period that is commonly used. Table 9 shows that the share of mispricing predictors which have a significantly negative long dummy and are thus classified as build-up predictors is 23.61% in my preferred specification (6) with return controls, year/month fixed effects and MHT controls. Moreover, 33% of mispricing predictors still have a significant positive long dummy, indicating that the mispricing associated with them has not been fully

Table 9: Share of Build-up & Unresolved Predictors - Errors 1 Year After Portfolio Formation



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	24869	24831	24831	24869	24831	24831
Number of Regressions	72	72	72	72	72	72

The table shows the result of the regression $forecast\, error_{i,d,t+12} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t}$, whereby $forecast\, error_{i,d,t+12}$ indicates the forecast error twelve months after portfolio formation. The regression is run for all predictors identified as mispricing in the regression shown in Column 3 of Table 3, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the return over the period between the formation of the predictor portfolio and the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

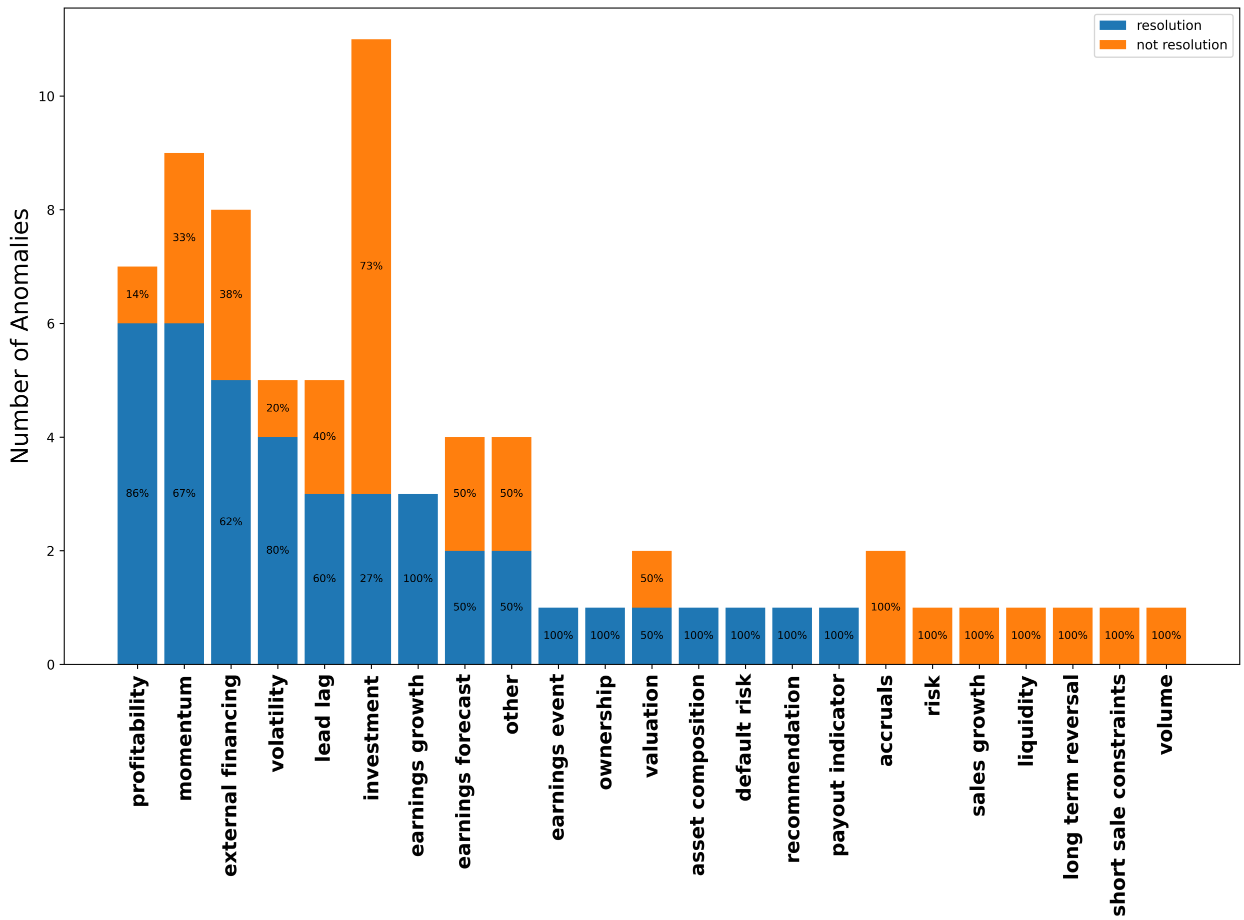
resolved.

To test if one year is a too short horizon for mispricing to build up, I repeat the analysis with forecasts made two years after portfolio formation (Table A4 in Appendix D) and find that almost no predictors can be classified as build-up. This suggests that the time over which mispricing is build up is generally less than two years.

4.2 Which Predictors are Build-up and Which are Resolution?

In this section, I separate the mispricing predictors by economic category using the classifications from A. Y. Chen and Zimmermann (2021) and examine each category's share of

Figure 3: Share of Resolution Predictors by Category



The figure shows the share of predictors classified as resolution by predictor category. The categories come from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification as resolution predictor is based on the regression of the forecast error at the time of portfolio formation on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Column 6 in Table 8).

resolution predictors. I use the regression specification with return controls, fixed effects and MHT controls (Column 6 in Table 8) for this classification.

Six out of seven mispricing predictors related to profitability are classified as resolution. This is in line with the results of Bouchaud et al. (2019) who show that the profitability predictor can be explained by sticky expectations that underreact to news. Similarly, the majority of momentum mispricings and all earnings growth-related mispricings are classified as resolution, both of which are commnly interpreted to reflect underreaction of the market to good news (Jegadeesh and Titman, 1993; Loh and Warachka, 2012).

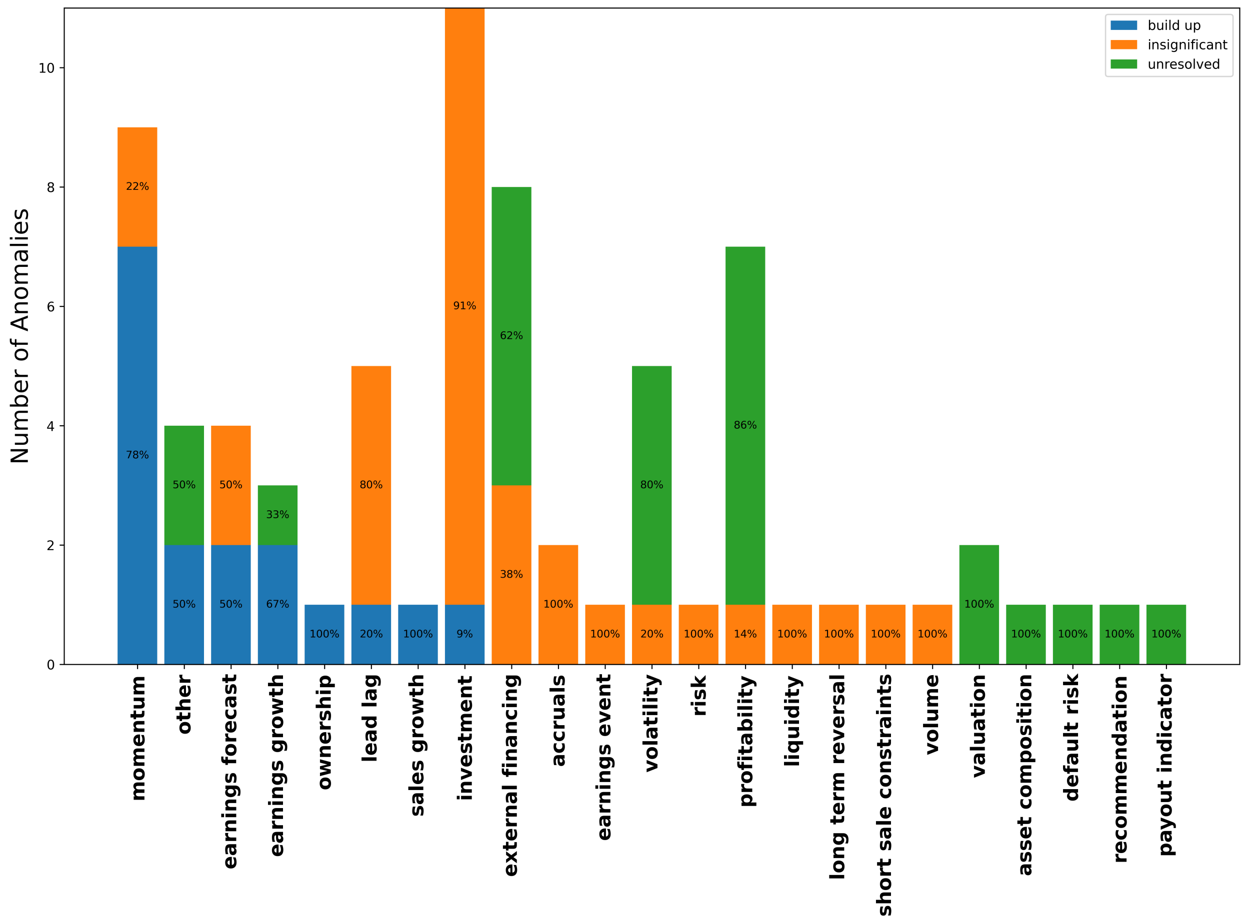
Moreover, a large share of external financing predictors is also classified as resolution. This is in line with Bradshaw et al. (2006) who show that high external financing is linked to analyst overoptimism.

Finally, four out of five volatility predictors are classified as resolution. Among those is the maximum return over the preceding month, which was originally also interpreted to capture the resolution of overpricing of stocks with high upsides (Bali et al., 2011) and analyst forecast dispersion, which was interpreted as resolving optimism driven by the fact that with high disagreement prices reflect the views of the more optimistic investors (Diether et al., 2002). The other volatility-related resolution predictors were originally not classified as mispricing.

Figure 4 shows the share of build-up and unresolved predictors by category. Interestingly, the majority of both momentum and earnings growth-related mispricings are classified as build-up in addition to being classified as resolution. This suggests that the market first underreacts to good news but then overcorrects for this initial underreaction, leading to a delayed overreaction. Both underreaction and delayed overreaction have been posed as explanations for momentum and my results suggest that both explanations have merit.

Moreover, a large share of the predictable mispricing related to external financing, volatility and profitability remains unresolved one year after portfolio formation, suggesting that mispricing from these sources is particularly long-lasting.

Figure 4: Share of Build-up and Unresolved Predictors by Category



The figure shows the share of predictors classified as build-up and unresolved by predictor category. The categories come from the Open Source Asset Pricing dataset (A. Y. Chen and Zimmermann, 2021). The classification is based on the regression of the forecast error one year after portfolio formation using forecasts for five quarters ahead on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Column 6 in Table 9).

5 Comparison to other Classification Methods

5.1 Mispricing Classification of the Original Authors

A. Y. Chen et al. (2022) hand-collect original authors' interpretation of whether the predictor they discovered represents mispricing or risk from the texts of the papers and find that "18% predictors are attributed to risk-based theory. 58% are attributed to mispricing and 24% have uncertain origins." They created a publicly available dataset including the predictor-level attributions, allowing me to compare them with my expectations-based classification.

I find that among the predictors classified as mispricing by the original authors, 46% are also classified as mispricing by my method. This relatively low share is not surprising given that I only detect a subset of mispricing types. Moreover, among the predictors classified as risk factors by the original authors, 40% are classified as mispricing by my method. This highlights that a substantial share of return predictability that was originally interpreted as being driven by risk is also associated with mispricing. However, as discussed in Section 2.2, my results do not allow me to determine if the return predictability is exclusively driven by mispricing or by a mixture of risk and mispricing.

Overall, the results show that the share of mispricings driven by biased expectations among supposed risk factors is relatively similar to that among supposed mispricings, which is in line with the results of A. Y. Chen et al. (2022), whose results also suggest that return predictors originally interpreted as risk are no more likely to be persistent risk factors than those originally interpreted as mispricing. Together, these results highlight the importance of formal classification methods in understanding the nature of a predictor.

5.2 Anomaly Classification of Holcblat et al. (2022)

Holcblat et al. (2022) classify predictors into anomalies and risk factors. They define an anomaly as any return predictor whose abnormal returns cannot be explained by risk, which is equivalent to my definition of mispricing. They perform their classification by testing if every risk-averse investor (defined as having a concave von Neumann-Morgenstern utility function) would prefer to invest in the predictor long over the predictor short portfolio. If they can reject this null hypothesis, at least some type of risk-averse investor would invest in the short portfolio. Therefore, its risk profile can potentially justify its lower returns. However, if the null hypothesis is true, no risk-averse investor would prefer the risk-return profile of the short portfolio over the long portfolio. Therefore, it must be mispriced.

While this pairwise portfolio test ignores the benefits of diversification that can be gained by investing in both the long and the short portfolio, Holcblat et al. (2022) also provide an equilibrium foundation, showing that their null hypothesis is not rejected if the return of a predictor exceeds the risk compensation required by risk-averse investors for a large class of utility functions.

Based on this argument, they classify a predictor as a **potential** risk factor if they reject the null hypothesis that every investor prefers the long over the short portfolio. There can still be mispricings among the potential risk factors because the test fails to classify a predictor as mispricing if a risk-averse utility function exists that rationalises a preference for the short over the long portfolio, regardless of whether any investor has such a utility function.

Any predictor for which the null hypothesis cannot be rejected is considered to be a mispricing predictor (anomaly). Hence, the mispricings detected by their test include predictors whose abnormal returns are driven by behavioural biases, information processing frictions, market frictions, and any other driver except risk.

Given that the classification of a predictor as mispricing is based on failing to reject a null hypothesis, a lack of statistical power can cause their test to mistakenly classify a risk factor as mispricing. They confirm with simulations that the test has good finite sample properties, so this is mostly a concern for return predictors with limited data availability. Nevertheless, this suggests that their test might be less suited to applications where falsely classifying a predictor as mispricing is more concerning than failing to detect a mispricing predictor, particularly for return predictors with relatively few available datapoints.

In contrast, my test focuses on biased expectations and thus only classifies return predictors that are driven by behavioural biases in information processing or information frictions as mispricing. On the one hand, this means that it provides a more specific classification. This is important because unlike mispricing from market frictions, mispricing due to biased expectations suggests trading strategies that yield abnormal returns and can appear out of sample through trading. On the other hand, it means that my test is less suited to accurately estimating the share of all predictors that is linked to mispricing and instead provides a lower bound on this share. Overall, our tests are complementary because they differ in the types of return predictors they classify as mispricing and because they have orthogonal reasons why they would fail to detect a mispricing.

Like me, Holcblat et al. (2022) apply their test to the predictor dataset from A. Y. Chen and Zimmermann (2021), allowing me to compare our classifications. I find that 65 of the 69 predictors that my test classifies as mispricing (using specification (6) of Table 3) are also mispricing according to their (unconditional) test, whereas only two are not classified as mispricing (*InvestPPEInv* & *TotalAccruals*) and two (*Recomm_ShortInterest* and *RealizedVol*) are missing from their data. However, they classify 86 additional predictors as mispricing. This result is not surprising since my test is more restrictive in the sources of return predictability it classifies as mispricing and statistical power issues cause my test to underestimate and their test to overestimate the share of mispricing predictors. In Table A9 in Appendix D, I show a predictor-level comparison between our tests for each predictor classified as mispricing by my test.

5.3 Build-up and Resolution Classification of van Binsbergen et al. (2023)

van Binsbergen et al. (2023) also classify predictors into build-up and resolution. Their method relies on measuring price wedges defined as the negative of the log of the fundamental value divided by the price. This means a negative price wedge implies underpricing, and a positive price wedge implies overpricing. The fundamental value is defined as the stream of future dividends discounted by a benchmark SDF. They estimate the price wedge at the portfolio level using 15 years of realised dividends. Their approach works for any factor-based SDF, and they use CAPM for their main result. A variable is considered to be a resolution predictor if the α and the long minus short price wedge have the opposite sign and a build-up predictor if they have the same sign. To determine the probability that a specific predictor portfolio is overpriced or underpriced, they use bootstrap simulations and study the fraction of simulated portfolios that are overpriced or underpriced. Because it is necessary to assume a specific SDF to apply the method, it is subject to the joint hypothesis problem. Thus, it may misclassify predictors if the wrong SDF is chosen.

van Binsbergen et al. (2023) do not provide a significance test for the long minus short price wedge, but they provide separate tests for the price wedges of the long and the short portfolio. I, therefore, consider a predictor to be a resolution (build-up) predictor if either the price wedge of the short portfolio is positive (negative) and significant or the price wedge of the long portfolio is negative (positive) and significant[10].

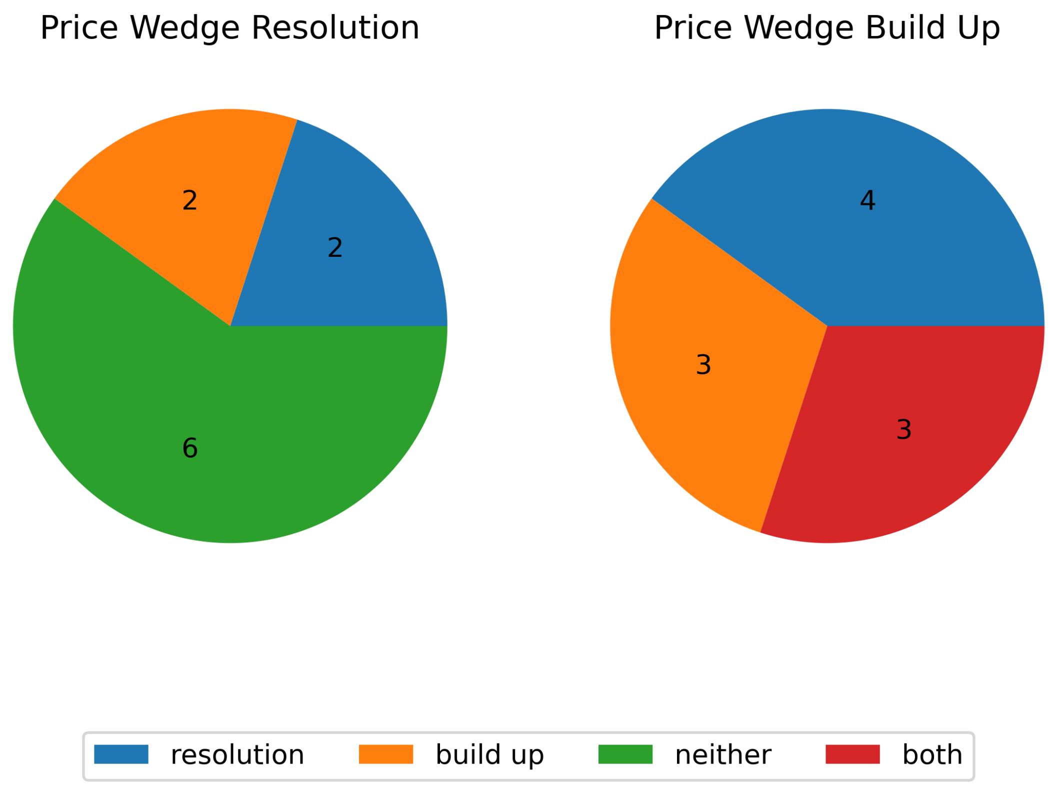
I cannot comprehensively compare our results as van Binsbergen et al. (2023) use a different set of predictors. However, there are 34 predictors from their sample for which I have the same or a closely related predictor in my sample. To keep coverage of as many variables as possible, I include variables from my sample even if they are not classified as mispricing in my first analysis step (Table 3. However, I still require a significant long minus short return using all returns up to the end of 2022, which reduces the overlapping set of predictors to 25.

In general, the expectation-based and price-wedge-based approaches to classify predictors into build-up and resolution are complementary because they use different data and have orthogonal strengths and weaknesses. However, Figure 5 shows that there is significant disagreement between the classification methods.

While it may seem like there is limited usefulness in combining both methods due to the large fraction of contradictions, this is potentially driven by the choice of CAPM as SDF. Future research using different SDFs could find less disagreement. Moreover, the comparison is based on just 25 predictors, and the overlap may be higher in a more complete sample of predictors.

[10]In theory it would be possible that a predictor is related to the build-up of mispricing in one leg of the sorted portfolio and the resolution of mispricing in the other leg, but in practice, this does not occur.

Figure 5: Comparison of expectation and price-wedge-based classifications



The figure shows the number of predictors classified as build-up/resolution by van Binsbergen et al. (2023) that is classified as build-up, resolution, neither or both by my method. The classification for my method is based on Column 6 of Table 8 for resolution predictors and Table 9 for build-up predictors.

6 Conclusion

Determining the degree of stock market efficiency is a fundamental question of finance with considerable implications for the efficiency of capital allocation and hence the real economy. Asset pricing researchers have discovered hundreds of characteristics that can predict the cross-section of stock returns, which challenges the efficient market hypothesis. However, this challenge crucially depends on whether these predictors represent mispricing or risk. My results suggest that for at least 40% of them, the predictable excess returns align with predictable changes in future earnings expectations, suggesting that they are at least partially driven by mispricing. This includes commonly used predictors such as momentum, profitability and investment. As my analysis does not capture mispricings stemming from biased beliefs about future required returns or market frictions and uses an imperfect proxy for the market's future earnings expectations, the actual share of mispricings among the predictors is likely even higher. Moreover, my results suggest that the excess returns of some predictors capture the build-up rather than the resolution of mispricing, implying that traders who capitalize on these predictors worsen rather than correct mispricing. Overall my results suggest that the ample evidence for predictability of the cross-section of returns does not just mean that existing asset pricing models have not yet incorporated all relevant risk factors but that there is widespread mispricing in the stock market.

References

Abarbanell, Jeffery S. and Brian J. Bushee (1998). "Abnormal Returns to a Fundamental Analysis Strategy". In: *The Accounting Review* 73.1, pp. 19–45. ISSN: 0001-4826.

Adam, Klaus and Stefan Nagel (2023). "Chapter 16 - Expectations Data in Asset Pricing". In: *Handbook of Economic Expectations*. Ed. by Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw. Academic Press, pp. 477–506. ISBN: 978-0-12-822927-9. DOI: 10.1016/B978-0-12-822927-9.00024-0.

Amihud, Yakov (2002). "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects". In: *Journal of Financial Markets* 5.1, pp. 31–56. ISSN: 1386-4181. DOI: 10.1016/S1386-4181(01)00024-6.

Anderson, Christopher W. and Luis Garcia-Feijóo (2006). "Empirical Evidence on Capital Investment, Growth Options, and Security Returns". In: *The Journal of Finance* 61.1, pp. 171–194. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2006.00833.x.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang (2006). "The Cross-Section of Volatility and Expected Returns". In: *The Journal of Finance* 61.1, pp. 259–299. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2006.00836.x.

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter (2005). "Short Interest, Institutional Ownership, and Stock Returns". In: *Journal of Financial Economics* 78.2, pp. 243–276. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2005.01.001.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov (2007). "Momentum and Credit Rating". In: *The Journal of Finance* 62.5, pp. 2503–2520. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2007.01282.x.

Balakrishnan, Karthik, Eli Bartov, and Lucile Faurel (2010). "Post Loss/Profit Announcement Drift". In: *Journal of Accounting and Economics* 50.1, pp. 20–41. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2009.12.002.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw (2011). "Maxing out: Stocks as Lotteries and the Cross-Section of Expected Returns". In: *Journal of Financial Economics* 99.2, pp. 427–446. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2010.08.014.

Banz, Rolf W. (1981). "The Relationship between Return and Market Value of Common Stocks". In: *Journal of Financial Economics* 9.1, pp. 3–18. ISSN: 0304-405X. DOI: 10.1016/0304-405X(81)90018-0.

Barbee, William C, Sandip Mukherji, and Gary A. Raines (1996). "Do Sales-Price and Debt-Equity Explain Stock Returns Better than Book-Market and Firm Size?" In: *Financial Analysts Journal* 52.2, pp. 56–60. ISSN: 0015-198X.

Bartov, Eli and Myungsun Kim (2004). "Risk, Mispricing, and Value Investing". In: *Review of Quantitative Finance and Accounting* 23.4, pp. 353–376. ISSN: 1573-7179. DOI: 10.1023/B:REQU.0000049321.34133.95.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch (2014). "Labor Hiring, Investment, and Stock Return Predictability in the Cross Section". In: *Journal of Political Economy* 122.1, pp. 129–177. ISSN: 0022-3808. DOI: 10.1086/674549.

Ben-Rephael, Azi, Steffen Hitzemann, and Yuanyuan Xiao (2021). *Investor (Mis)Reaction, Biased Beliefs, and the Mispricing Cycle*. SSRN Scholarly Paper ID 3966758. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3966758.

Benjamini, Yoav and Yosef Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1, pp. 289–300. ISSN: 0035-9246.

Bhandari, Laxmi Chand (1988). "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence". In: *The Journal of Finance* 43.2, pp. 507–528. ISSN: 0022-1082. DOI: 10.2307/2328473.

Van Binsbergen, Jules H., Martijn Boons, Christian C. Opp, and Andrea Tamoni (2023). "Dynamic Asset (Mis)Pricing: Build-up versus Resolution Anomalies". In: *Journal of Financial Economics* 147.2, pp. 406–431. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2022.11.005.

Van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira (2022). *Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases*. DOI: 10.2139/ssrn.3625279. preprint.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic Expectations and Stock Returns". In: *The Journal of Finance* 74.6, pp. 2839–2874. ISSN: 1540-6261. DOI: 10.1111/jofi.12833.

Bordalo, Pedro, Nicola Gennaioli, Rafael Laporta, and Andrei Shleifer (2023). *Belief Overreaction and Stock Market Puzzles*. URL: https://papers.ssrn.com/abstract=3615480 (visited on 10/19/2023). preprint.

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier, and David Thesmar (2019). "Sticky Expectations and the Profitability Anomaly". In: *The Journal of Finance* 74.2, pp. 639–674. ISSN: 1540-6261. DOI: 10.1111/jofi.12734.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts (2007). "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing". In: *The Journal of Finance* 62.2, pp. 877–915. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2007.01226.x.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan (2006). "The Relation between Corporate Financing Activities, Analysts' Forecasts and Stock Returns". In: *Journal of Accounting and Economics*. Conference Issue on Implications of Changing Financial Reporting Standards 42.1, pp. 53–85. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2006.03.004.

Campbell, John Y. (1991). "A Variance Decomposition for Stock Returns". In: *The Economic Journal* 101.405, pp. 157–179. ISSN: 0013-0133. DOI: 10.2307/2233809.

Campbell, John Y. and Robert J. Shiller (1988a). "Stock Prices, Earnings, and Expected Dividends". In: *The Journal of Finance* 43.3, pp. 661–676. ISSN: 0022-1082. DOI: 10.2307/2328190.

— (1988b). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". In: *The Review of Financial Studies* 1.3, pp. 195–228. ISSN: 0893-9454.

Carhart, Mark M. (1997). "On Persistence in Mutual Fund Performance". In: *The Journal of Finance* 52.1, pp. 57–82. ISSN: 0022-1082. DOI: 10.2307/2329556.

Cen, Ling, K C John Wei, and Jie Zhang (2006). "FEPS - Forecasted Earnings per Share and the Cross Section of Expected Stock Returns". In: p. 47.

Chan, Louis K. C., Narasimhan Jegadeesh, and Josef Lakonishok (1996). "Momentum Strategies". In: *The Journal of Finance* 51.5, pp. 1681–1713. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1996.tb05222.x.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis (2001). "The Stock Market Valuation of Research and Development Expenditures". In: *The Journal of Finance* 56.6, pp. 2431–2456. ISSN: 1540-6261. DOI: 10.1111/0022-1082.00411.

Charoenrook, Anchada Aida and Jennifer S. Conrad (2005). *Identifying Risk-Based Factors*. DOI: 10.2139/ssrn.757210. URL: https://papers.ssrn.com/abstract=757210 (visited on 09/15/2022). preprint.

Chen, Andrew Y., Alejandro Lopez-Lira, and Tom Zimmermann (2022). *Peer-Reviewed Theory Does Not Help Predict the Cross-section of Stock Returns*. DOI: 10.2139/ssrn.4308069. URL: https://papers.ssrn.com/abstract=4308069 (visited on 01/20/2023). preprint.

Chen, Andrew Y. and Tom Zimmermann (2021). *Open Source Cross-Sectional Asset Pricing*. SSRN Scholarly Paper 3604626. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3604626.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein (2002). "Breadth of Ownership and Stock Returns". In: *Journal of Financial Economics*. Limits on Arbitrage 66.2, pp. 171–205. ISSN: 0304-405X. DOI: 10.1016/S0304-405X(02)00223-4.

Chen, Long, Zhi Da, and Xinlei Zhao (2013). "What Drives Stock Price Movements?" In: *Review of Financial Studies* 26.4, pp. 841–876. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hht005.

Chordia, Tarun, Amit Goyal, and Alessio Saretto (2020). "Anomalies and False Rejections". In: *The Review of Financial Studies* 33.5, pp. 2134–2179. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa018.

Cohen, Lauren and Dong Lou (2012). "Complicated Firms". In: *Journal of Financial Economics*. Special Issue on Investor Sentiment 104.2, pp. 383–400. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2011.08.006.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill (2008). "Asset Growth and the Cross-Section of Stock Returns". In: *The Journal of Finance* 63.4, pp. 1609–1651. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2008.01370.x.

Cuevas Rodriguez, Gabriel, Denis Mokanov, and Danyu (Jane) Zhang (2023). *Earnings Expectations and Asset Prices*. DOI: 10.2139/ssrn.4572199. URL: https://papers.ssrn.com/abstract=4572199 (visited on 11/07/2023). preprint.

Da, Zhi and Mitch Warachka (2011). "EarningsForecastDisparity - The Disparity between Long-Term and Short-Term Forecasted Earnings Growth". In: *Journal of Financial Economics* 100.2, pp. 424–442. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2010.10.015.

Daniel, Kent and Sheridan Titman (2006). "Market Reactions to Tangible and Intangible Information". In: *The Journal of Finance* 61.4, pp. 1605–1643. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2006.00884.x.

De Bondt, Werner F. M. and Richard Thaler (1985). "Does the Stock Market Overreact?" In: *The Journal of Finance* 40.3, pp. 793–805. ISSN: 0022-1082. DOI: 10.2307/2327804.

De La O, Ricardo and Sean Myers (2021). "Subjective Cash Flow and Discount Rate Expectations". In: *The Journal of Finance* 76.3, pp. 1339–1387. ISSN: 1540-6261. DOI: 10.1111/jofi.13016.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina (2002). "ForecastDispersion - Differences of Opinion and the Cross Section of Stock Returns". In: *The Journal of Finance* 57.5, pp. 2113–2141. ISSN: 1540-6261. DOI: 10.1111/0022-1082.00490.

Drake, Michael S., Lynn Rees, and Edward P. Swanson (2011). "Should Investors Follow the Prophets or the Bears? Evidence on the Use of Public Information by Analysts and Short Sellers". In: *The Accounting Review* 86.1, pp. 101–130. ISSN: 0001-4826.

Engelberg, Joseph, R. David McLean, and Jeffrey Pontiff (2020). "Analysts and Anomalies". In: *Journal of Accounting and Economics* 69.1, p. 101249. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2019.101249.

Engelberg, Joseph, R. David Mclean, and Jeffrey Pontiff (2018). "Anomalies and News". In: *The Journal of Finance* 73.5, pp. 1971–2001. ISSN: 1540-6261. DOI: 10.1111/jofi.12718.

Fama, Eugene F. and Kenneth R. French (1992). "The Cross-Section of Expected Stock Returns". In: *The Journal of Finance* 47.2, pp. 427–465. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1992.tb04398.x.

— (2006). "Profitability, Investment and Average Returns". In: *Journal of Financial Economics* 82.3, pp. 491–518. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2005.09.009.

— (2015). "A Five-Factor Asset Pricing Model". In: *Journal of Financial Economics* 116.1, pp. 1–22. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2014.10.010.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu (2020). "Taming the Factor Zoo: A Test of New Factors". In: *The Journal of Finance* 75.3, pp. 1327–1370. ISSN: 1540-6261. DOI: 10.1111/jofi.12883.

Gafka, Beata, Pavel G. Savor, and Mungo Ivor Wilson (2021). *Sources of Return Predictability*. DOI: 10.2139/ssrn.3937343. URL: https://papers.ssrn.com/abstract=3937343 (visited on 06/08/2023). preprint.

Giglio, Stefano and Bryan Kelly (2018). "Excess Volatility: Beyond Discount Rates". In: *The Quarterly Journal of Economics* 133.1, pp. 71–127. ISSN: 0033-5533. DOI: 10.1093/qje/qjx034.

Giglio, Stefano, Yuan Liao, and Dacheng Xiu (2021). "Thousands of Alpha Tests". In: *The Review of Financial Studies* 34.7, pp. 3456–3496. ISSN: 0893-9454. DOI: 10.1093/rfs/hhaa111.

Grinblatt, Mark, Gergana Jostova, and Alexander Philipov (2018). "Analyst Bias and Mispricing". In: p. 49.

Guo, Re-Jin, Baruch Lev, and Charles Shi (2006). "Explaining the Short- and Long-Term IPO Anomalies in the US by R&D". In: *Journal of Business Finance & Accounting* 33.3-4, pp. 550–579. ISSN: 1468-5957. DOI: 10.1111/j.1468-5957.2006.00610.x.

Guo, Li, Frank Weikai Li, and K. C. John Wei (2020). "Security Analysts and Capital Market Anomalies". In: *Journal of Financial Economics* 137.1, pp. 204–230. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2020.01.002.

Hafzalla, Nader, Russell Lundholm, and E. Matthew Van Winkle (2011). "Percent Accruals". In: *The Accounting Review* 86.1, pp. 209–236. ISSN: 0001-4826.

Han, Yufeng, Yueliang (Jacques) Lu, Weike Xu, and Guofu Zhou (2020). *Mispricing and Anomalies: An Exogenous Shock to Short Selling from JGTRRA*. DOI: 10.2139/ssrn.3730653. URL: https://papers.ssrn.com/abstract=3730653 (visited on 11/07/2023). preprint.

Han, Yufeng, Guofu Zhou, and Yingzi Zhu (2016). "A Trend Factor: Any Economic Gains from Using Information over Investment Horizons?" In: *Journal of Financial Economics* 122.2, pp. 352–375. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2016.01.029.

Harvey, Campbell R. and Yan Liu (2021). "Lucky Factors". In: *Journal of Financial Economics* 141.2, pp. 413–435. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2021.04.014.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (2016). "… and the Cross-Section of Expected Returns". In: *The Review of Financial Studies* 29.1, pp. 5–68. ISSN: 0893-9454. DOI: 10.1093/rfs/hhv059.

Hawkins, Eugene H., Stanley C. Chamberlin, and Wayne E. Daniel (1984). "AnalystRevision - Earnings Expectations and Security Prices". In: *Financial Analysts Journal* 40.5, pp. 24–74. ISSN: 0015-198X.

He, Ai, Dashan Huang, Jiaen Li, and Guofu Zhou (2022). "Shrinking Factor Dimension: A Reduced-Rank Approach". In: *Management Science*. ISSN: 0025-1909. DOI: 10.1287/mnsc.2022.4563.

Heston, Steven L. and Ronnie Sadka (2008). "Seasonality in the Cross-Section of Stock Returns". In: *Journal of Financial Economics* 87.2, pp. 418–445. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2007.02.003.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang (2004). "Do Investors Overvalue Firms with Bloated Balance Sheets?" In: *Journal of Accounting and Economics*. Conference Issue on Research on Market Efficiency, Valuation, and Mispricing: Risk, Behavioral, and Alternative Explanations 38, pp. 297–331. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2004.10.002.

Holcblat, Benjamin, Abraham Lioui, and Michael Weber (2022). *Anomaly or Possible Risk Factor? Simple-To-Use Tests*. DOI: 10.2139/ssrn.4064159. URL: https://papers.ssrn.com/abstract=4064159 (visited on 07/05/2022). preprint.

Hong, Harrison and Marcin Kacperczyk (2009). "The Price of Sin: The Effects of Social Norms on Markets". In: *Journal of Financial Economics* 93.1, pp. 15–36. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2008.09.001.

Hou, Kewei (2007). "Industry Information Diffusion and the Lead-lag Effect in Stock Returns". In: *The Review of Financial Studies* 20.4, pp. 1113–1138. ISSN: 0893-9454. DOI: 10.1093/revfin/hhm003.

Hou, Kewei and Tobias J. Moskowitz (2005). "Market Frictions, Price Delay, and the Cross-Section of Expected Returns". In: *The Review of Financial Studies* 18.3, pp. 981–1020. ISSN: 0893-9454. DOI: 10.1093/rfs/hhi023.

Hou, Kewei, Chen Xue, and Lu Zhang (2020). "Replicating Anomalies". In: *The Review of Financial Studies* 33.5, pp. 2019–2133. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy131.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen (1995). "Market Underreaction to Open Market Share Repurchases". In: *Journal of Financial Economics* 39.2, pp. 181–208. ISSN: 0304-405X. DOI: 10.1016/0304-405X(95)00826-Z.

Jackson, Andrew and Timothy Johnson (2006). "Unifying Underreaction Anomalies". In: *The Journal of Business* 79.1, pp. 75–114. ISSN: 0021-9398. DOI: 10.1086/497406.

Jegadeesh, Narasimhan (1990). "Evidence of Predictable Behavior of Security Returns". In: *The Journal of Finance* 45.3, pp. 881–898. ISSN: 0022-1082. DOI: 10.2307/2328797.

Jegadeesh, Narasimhan and Sheridan Titman (1993). "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency". In: *The Journal of Finance* 48.1, pp. 65–91. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1993.tb04702.x.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen (2023). "Is There a Replication Crisis in Finance?" In: *The Journal of Finance* n/a.n/a. ISSN: 1540-6261. DOI: 10.1111/jofi.13249.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh (2018). "Interpreting Factor Models". In: *The Journal of Finance* 73.3, pp. 1183–1223. ISSN: 1540-6261. DOI: 10.1111/jofi.12612.

La Porta, Rafael (1996). "Expectations and the Cross-Section of Stock Returns". In: *The Journal of Finance* 51.5, pp. 1715–1742. ISSN: 0022-1082. DOI: 10.2307/2329535.

La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny (1997). "Good News for Value Stocks: Further Evidence on Market Efficiency". In: *The Journal of Finance* 52.2, pp. 859–874. ISSN: 0022-1082. DOI: 10.2307/2329502.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny (1994). "Contrarian Investment, Extrapolation, and Risk". In: *The Journal of Finance* 49.5, pp. 1541–1578. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1994.tb04772.x.

Landsman, Wayne R., Bruce L. Miller, Ken Peasnell, and Shu Yeh (2011). "Do Investors Understand Really Dirty Surplus?" In: *The Accounting Review* 86.1, pp. 237–258. ISSN: 0001-4826.

Lee, Charles M.C. and Bhaskaran Swaminathan (2000). "Price Momentum and Trading Volume". In: *The Journal of Finance* 55.5, pp. 2017–2069. ISSN: 1540-6261. DOI: 10.1111/0022-1082.00280.

Lev, Baruch and Doron Nissim (2004). "Taxable Income, Future Earnings, and Equity Values". In: *The Accounting Review* 79.4, pp. 1039–1074. ISSN: 0001-4826.

Lintner, John (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets". In: *The Review of Economics and Statistics* 47.1, pp. 13–37. ISSN: 0034-6535. DOI: 10.2307/1924119.

Liu, Weimin (2006). "A Liquidity-Augmented Capital Asset Pricing Model". In: *Journal of Financial Economics* 82.3, pp. 631–671. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2005.10.001.

Lockwood, Larry and Wikrom Prombutr (2010). "Sustainable Growth and Stock Returns". In: *Journal of Financial Research* 33.4, pp. 519–538. ISSN: 1475-6803. DOI: 10.1111/j.1475-6803.2010.01281.x.

Loh, Roger K. and Mitch Warachka (2012). "Streaks in Earnings Surprises and the Cross-Section of Stock Returns". In: *Management Science* 58.7, pp. 1305–1321. ISSN: 0025-1909. DOI: 10.1287/mnsc.1110.1485.

Lyandres, Evgeny, Le Sun, and Lu Zhang (2008). "The New Issues Puzzle: Testing the Investment-Based Explanation". In: *The Review of Financial Studies* 21.6, pp. 2825–2855. ISSN: 0893-9454. DOI: 10.1093/rfs/hhm058.

Nguyen, Giao X. and Peggy E. Swanson (2009). "Firm Characteristics, Relative Efficiency, and Equity Returns". In: *Journal of Financial and Quantitative Analysis* 44.1, pp. 213–236. ISSN: 0022-1090, 1756-6916. DOI: 10.1017/S0022109009090012.

Novy-Marx, Robert (2012). "Is Momentum Really Momentum?" In: *Journal of Financial Economics* 103.3, pp. 429–453. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2011.05.003.

Pástor, Ľuboš and Robert F. Stambaugh (2003). "Liquidity Risk and Expected Stock Returns". In: *Journal of Political Economy* 111.3, pp. 642–685. ISSN: 0022-3808. DOI: 10.1086/374184.

Pontiff, Jeffrey and Artemiza Woodgate (2008). "Share Issuance and Cross-sectional Returns". In: *The Journal of Finance* 63.2, pp. 921–945. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2008.01335.x.

Pruitt, Seth (2023). *Dogs and Cats Living Together: A Defense of Cash-Flow Predictability.* DOI: 10.2139/ssrn.4343101. URL: https://papers.ssrn.com/abstract=4343101 (visited on 06/27/2023). preprint.

Pukthuanthong, Kuntara, Richard Roll, and Avanidhar Subrahmanyam (2019). "A Protocol for Factor Identification". In: *The Review of Financial Studies* 32.4, pp. 1573–1607. ISSN: 0893-9454. DOI: 10.1093/rfs/hhy093.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman, and İrem Tuna (2005). "Accrual Reliability, Earnings Persistence and Stock Prices". In: *Journal of Accounting and Economics* 39.3, pp. 437–485. ISSN: 0165-4101. DOI: 10.1016/j.jacceco.2005.04.005.

Sharpe, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*". In: *The Journal of Finance* 19.3, pp. 425–442. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1964.tb02865.x.

Sloan, Richard G. (1996). "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?" In: *The Accounting Review* 71.3.

Thomas, Jacob K. and Huai Zhang (2002). "Inventory Changes and Future Returns". In: *Review of Accounting Studies* 7.2, pp. 163–187. ISSN: 1573-7136. DOI: 10.1023/A:1020221918065.

Xie, Hong (2001). "The Mispricing of Abnormal Accruals". In: *The Accounting Review* 76.3, pp. 357–373. ISSN: 0001-4826.

Yan, Shu (2011). "Jump Risk, Stock Returns, and Slope of Implied Volatility Smile". In: *Journal of Financial Economics* 99.1, pp. 216–233. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2010.08.011.

Zhang, X. Frank (2006). "Information Uncertainty and Stock Returns". In: *The Journal of Finance* 61.1, pp. 105–137. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.2006.00831.x.

A Derivations of Equation 1 & 2

The log return of an asset between time t and time $t+1$ is given by the following equation, where P_t is the assets price at time t and D_{t+1} is the dividend paid between time t and $t+1$:

$$r_{t+1} = log(P_{t+1} + D_{t+1}) - log(P_t)) \tag{8}$$

Based on this, Campbell and Shiller (1988a,b) showed that the following linear relationship between the log return, the log dividend and the log price holds approximately:

$$r_{t+1} = k + \rho p_{t+1} + (1-\rho)d_{t+1} - p_t \tag{9}$$

In this equation, d_t is the log dividend paid at time t, and p_t is the log stock price. The parameter ρ is related to the price-dividend ratio and is close to but smaller than one, and k is a constant term. Each variable can refer to an individual stock or the average in a portfolio of stocks. Following Campbell (1991) this equation can be rewritten as follows:

$$r_{t+1} = k - \rho(d_{t+1} - p_{t+1}) + d_{t+1} - p_t \tag{10}$$

Defining $g_{t+1} = d_{t+1} - d_t$ and rearranging gives:

$$p_t - d_t = k + \rho(p_{t+1} - d_{t+1}) - r_{t+1} + g_{t+1} \tag{11}$$

Iterating the equation forward while imposing $\lim\limits_{s\to\infty} \rho^s(d_{t+s} - p_{t+s}) = 0$ gives:

$$p_t - d_t = \frac{k}{1-\rho} + \sum_{s=0}^{\infty} \rho^s g_{t+1+s} - \sum_{s=0}^{\infty} \rho^s r_{t+1+s} \tag{12}$$

Bringing $p_t - d_t$ to the right side and r_{t+1} to the left side gives:

$$r_{t+1} = \frac{k}{1-\rho} + \sum_{s=0}^{\infty} \rho^s g_{t+s} - \sum_{s=1}^{\infty} \rho^s r_{t+s} - (p_t - d_t) \tag{13}$$

This holds in expectation

$$\mathbb{E}_t(r_{t+1}) = \frac{k}{1-\rho} + \mathbb{E}_t(\sum_{s=0}^{\infty} \rho^s g_{t+s}) - \mathbb{E}_t(\sum_{s=1}^{\infty} \rho^s r_{t+s}) + (d_t - p_t) \tag{14}$$

The realised return for a period is then the expected return plus the change in dividend expectations and minus the change in return expectations:

$$r_{t+1} = E_t(r_{t+1}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} \rho^s g_{t+1+s} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} \rho^s r_{t+1+s} \tag{15}$$

Campbell and Shiller (1988a,b) originally used the log linearisation to study the market portfolio. However, since it follows directly from the definition of the log return, it can be applied to other portfolios and individual stocks, as I do in Equation 1.

If it is applied to a portfolio, the (potentially weighted) average dividend and price of the portfolio need to be calculated before doing the log transformation since the average of a log is not the same as the log of the average. Notably, it is theoretically possible to aggregate the long-minus-short portfolio into a single asset by taking the average price (dividend) of the short portfolio and subtracting it from that of the long portfolio before applying the log-linearisation. However, this is impractical because it is not guaranteed that the resulting averages would all be positive, which is necessary to take the log. Therefore, I average over the long and short portfolios separately and define:

$$d_t^L = log(\sum_{i \in L_t}(w_i \cdot D_{i,t})) \tag{16}$$

$$P_t^L = log(\sum_{i \in L_t}(w_i \cdot P_{i,t})) \tag{17}$$

In the equation above, L_t is the set of all stocks that are in the long portfolio at time t and w_i is the weight applied to a stock. Common weighting methods are equal weighting and value weighting. Applying the Cambell-Shiller log-linearisation and solving forward, as shown above, we can then write:

$$r_{t+1}^L = E_t(r_{t+1}^L) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^L)^s g_{t+1+s}^L - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^L)^s r_{t+1+s}^L \tag{18}$$

We can define expressions for S_t, the set of all stocks in the short portfolio at time t accordingly. We can then write:

$$r_t^{LS} = r_t^L - r_t^S \tag{19}$$

$$g_t^{LS} = g_t^L - g_t^S \tag{20}$$

This gives us the equation for the the long-minus-short return (Equation 2):

$$r_{t+1}^{LS} \approx \mathbb{E}_t(r_{t+1}^{LS}) + [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=0}^{\infty} (\rho^{LS})^s g_{t+1+s}^{LS} - [\mathbb{E}_{t+1} - \mathbb{E}_t] \sum_{s=1}^{\infty} (\rho^{LS})^s r_{t+1+s}^{LS} \tag{21}$$

This equation requires a further approximation because ρ^L and ρ^S are not exactly the same. However, they are both close to 1 by construction, so we can approximate them both by ρ^{LS}.

B Magnitude

In this appendix, I provide a back-of-the-envelope calculation to estimate the magnitude of the return differential between the long and short portfolios of mispricing predictors that the predictable differences in forecast revisions can explain.

From Equation 1, we know that changes in expectations for all future dividends between time t and time $t + 1$ are a central driver of the return between time t and time $t + 1$. Throughout this paper, I use quarterly earnings forecasts as a proxy for dividend expectations. Therefore, to estimate if changes in expectations can justify the higher return of stocks in the long portfolio, I will estimate the monthly changes in earnings expectations for every firm in the long and short portfolio and calculate the return implied by these changes in expectations. By comparing the difference in implied returns of stocks in the long portfolio and stocks in the short portfolio to the actually observed return difference between the portfolios, I can then asses how much of the observed return differential can be explained by predictable differences in expectation revisions. I do this separately for all 68 predictors that I identified as mispricing in the analysis in Section 3.

Estimating the implied returns faces a significant issue: it requires a forecast for all future earnings over a firm's life cycle. However, my dataset only contains forecasts for up to eight quarters ahead and a long-term growth forecast, representing the yearly growth rate over the next business cycle of three to five years. Moreover, all eight quarterly forecasts and the LTG forecasts are rarely available for a given firm in a given month, and often, not a single forecast is available. Therefore, to calculate the implied returns, I need to impute the missing quarterly forecasts for up to eight quarters ahead and then calculate the estimated forecasts for quarters more than eight quarters into the future.

This heavy use of imputing implies that my estimates will be rough and the exact numbers should not be over-interpreted. However, this method allows me to asses if predictable differences in earnings revisions between stocks in the long portfolio are sufficiently large to explain the entire predictable differences in returns.

To calculate the implied return of the long and short portfolio, I proceed as follows: On the first day of each month, I calculate each firm's median earnings forecast for one to eight quarters ahead and the median long-term growth forecast. I take forecasts from individual analysts made up to 28 days before the beginning of the month into account for the calculation of the median. Then, I retain all firms for which I have an available LTG forecast and at least one quarterly forecast. Next, I impute missing quarterly forecasts using the forecast for the preceding quarter and the forecasted quarterly earnings growth implied by the LTG forecast. Thus, I impute a missing forecast for firm i at month t for q quarters ahead for $q \in \{2...8\}$ as:

$$forecast_{i,t,q} = forecast_{i,t,q-1} * (1 + LTG)^{\frac{1}{4}} \tag{22}$$

If $forecast_{i,t,q-1}$ is negative, I do not use it for imputation because long-term growth rates are not meant to be applied to negative forecasts, as a positive LTG forecast always reflects a belief that the firm will grow.

Next, I perform the same procedure backwards; that is, I use available forecasts for the subsequent quarter to impute forecasts that are still missing after the first step of imputation for $q \in \{1...7\}$:

$$forecast_{i,t,q} = forecast_{i,t,q+1}/(1 + LTG)^{\frac{1}{4}} \tag{23}$$

Afterwards, I drop all forecasts for which the eight quarters ahead forecast could not be imputed and then iteratively calculate the forecasts for nine quarters ahead up to 20 quarters ahead, using Equation 22 for $q \in \{9...20\}$.

Since LTG only captures expectations for up to five years ahead, it is not a measure of growth expectations for quarters beyond 20 quarters ahead. Therefore, I calculate the forecasts for between 20 and 200 quarters ahead, assuming a constant yearly growth rate of 4%. Next, I calculate for each month the average forecast in the long and short portfolio across stocks for between 1 and 200 quarters ahead, as well as the average stock price in each portfolio.

I assume that all firms stop existing after 200 quarters, which is a mild assumption since, at this point, future earnings are discounted so heavily that they have minimal effect. Furthermore, I assume that the discount factor (r) is constant over time. Under these assumptions, the price of a stock is given by:

$$P_t = \sum_{q=1}^{200} \frac{\tilde{\mathbb{E}}_t(e_{t+q})}{\tilde{\mathbb{E}}_t(r^q)} \tag{24}$$

Using the average forecasts for up to 200 quarters ahead and the average price, I can numerically solve Equation 24 for the discount factor (r) that justifies the currently observed average stock price, given the forecasted future earnings.

Then, I calculate the one-month-ahead price (P_{t+1}) that is implied by the revised earnings forecasts one month later by substituting the revised forecasts for between 1 and 200 quarters ahead and the discount factor (r) into equation Equation 24. This allows me to calculate the implied monthly return by comparing P_{t+1} and P_t.

I separately calculate the implied monthly returns for stocks in the long and short portfolios in each month and then take the average across months in each portfolio. This gives me the average monthly return of the long and the short portfolio implied by the monthly revisions in earnings forecasts. Next, I calculate the implied long-minus short return by subtracting the implied short return from the implied long return.

As a final step, I calculate the actual difference in return between stocks in the long portfolio, using only the subset of stocks with available forecasts that were also included in the calculation of the implied return difference.

For a few predictors, I have several months where I do not have a single firm with a valid data point in one of the portfolios. As this is a sign of generally bad data availability and suggests that my results only reflect a small percentage of the total stocks in the long and short portfolio, I drop predictors where this is the case in more than ten months from my analysis. These predictors often also have unreasonably extreme implied returns, indicating that outliers drive them. This leaves me with 44 out of the 68 mispricing predictors.

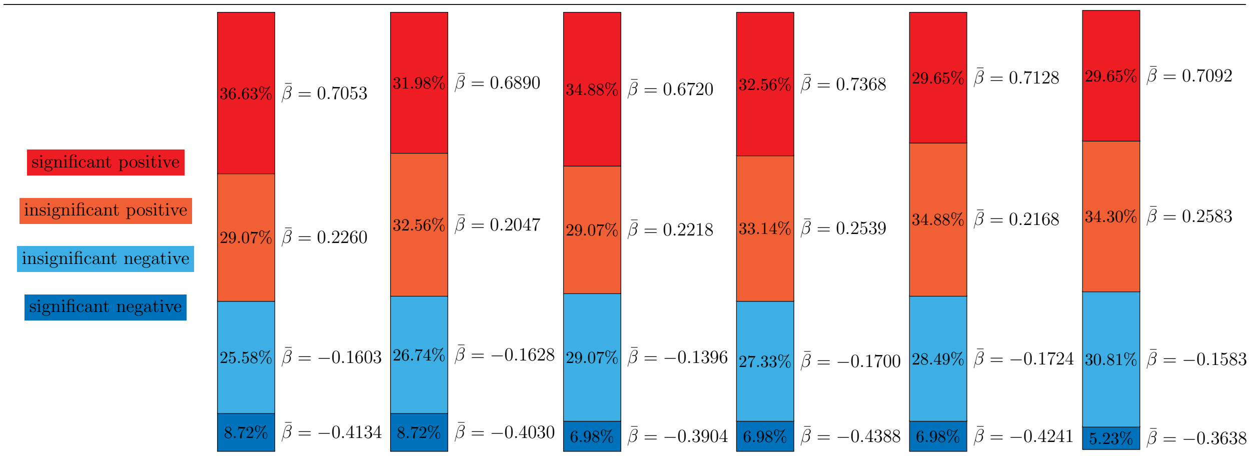
I find that, averaged across the 44 predictors, the monthly long-minus-short return that is implied by the earnings forecast revisions is 0.70%, whereas the actual return difference is 0.35%. As discussed above, this is a rough estimate, and the result depends heavily on the LTG forecasts and the quarterly forecast with the longest available horizon, as those forecasts are used in the calculation of the quarterly forecasts with horizons beyond eight quarters ahead, which drive a large share of the implied return. Moreover, I need to assume a constant discount factor and use the observed stock price, which does not fully reflect the fundamental value of the stocks, given that they are part of mispricing predictor portfolios.

However, despite the inaccuracy of my estimation procedure, it suggests that the implied returns are in the correct order of magnitude to fully explain the observed return difference.

C Long-Term Growth Data

Table A1: Share of Mispricings among the Predictors - Long-Term Growth Data

	(1)	(2)	(3)	(4)	(5)	(6)



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	36129	36129	36129	36129	36129	36129
Number of Regressions	172	172	172	172	172	172

The table shows the result of the regression $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision_{i,d,t,x}$ is the revision of the consensus long-term earnings growth forecast. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

Table A1 shows the same regression analysis as Table 3, except that it uses long-term growth (LTG) forecasts from IBES, whereby LTG is defined as the yearly earnings growth rate over a complete business cycle of three to five years. The calculation of the LTG forecast revision follows the one for the quarterly forecast revision (Equation 3), except that I do not scale by price because the forecast is for a growth rate and not a level. Moreover, since LTG forecasts have no fixed target date, keeping the target date constant is impossible. Hence, the first and the second forecasts may describe earnings growth rates over slightly different periods.

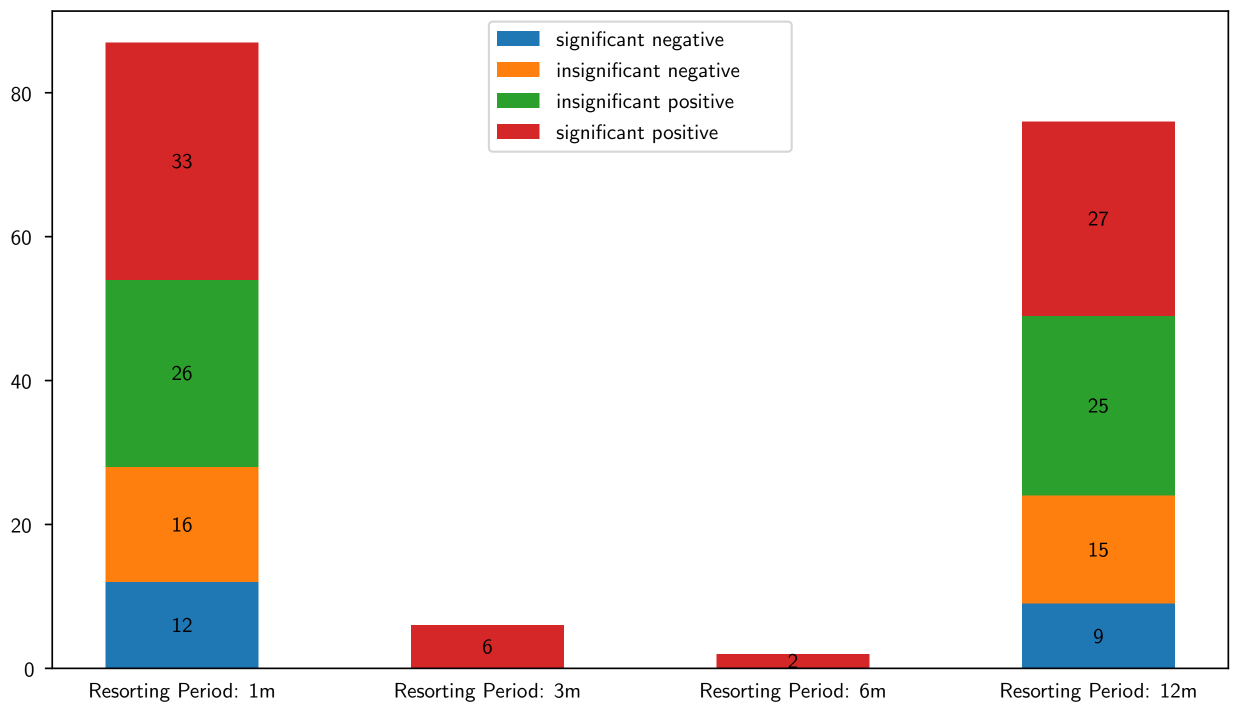
The share of predictors with a positive coefficient is similar for quarterly and LTG forecasts. However, slightly fewer results are significant with LTG forecasts, likely due to worse data availability. Next, I study if the same predictors are classified as mispricing with both types of forecasts, focusing on the specification with return controls, fixed effects and MHT controls. Thirty-one predictors have positive significant coefficients with both LTG and quarterly data. Twenty additional predictors have significant coefficients with LTG data.

Moreover, 37 predictors have significant positive coefficients with quarterly forecasts, and three of those have a significant negative long dummy using LTG forecast.

The fact that some predictors are only significant for one of the two forecast horizons and that there are even a few cases with significant coefficients that go in opposite directions is not a challenge to the validity of my methodology. As Equation 1 shows, returns are not related to expectations for a specific horizon but to expectations about the sum of all future dividends. Therefore, a mispricing predictor may predict relatively higher returns and relatively more negative forecast revisions for a forecast horizon if it predicts more positive forecast revisions for other horizons. In contrast, risk factors can never predict earnings forecast revisions of any kind. Therefore, if anything, this result implies that even more predictors are linked to mispricing than my analysis in Section 3 suggests since a significant link between the portfolio assignment and either the quarterly or the LTG forecast revision is sufficient to establish that expectations are biased.

D Additional Figures & Tables

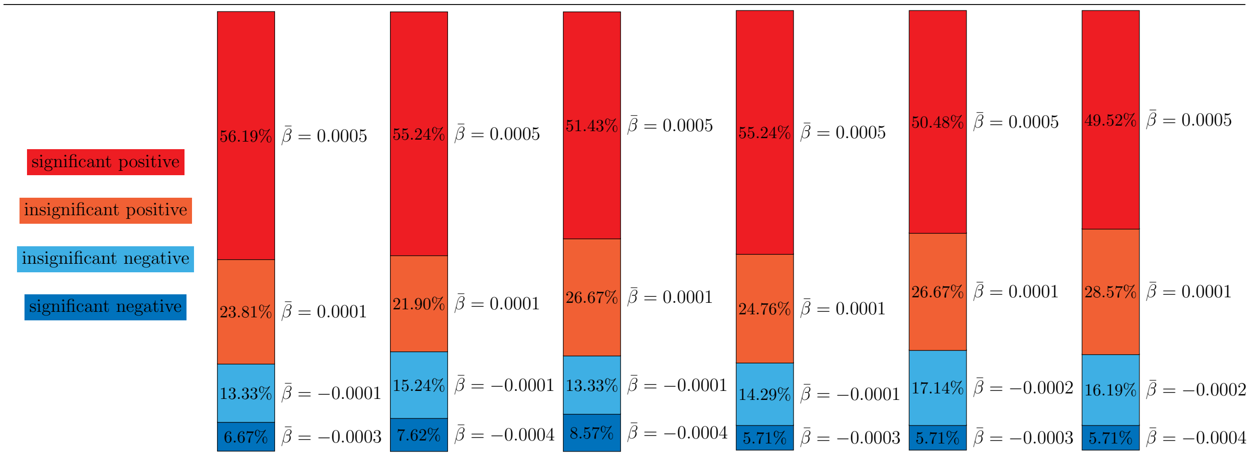
Figure A1: Share of Mispricings by Resorting Period



The figure shows the share of predictors that are classified as mispricing by resorting period. The classification as mispricing is based on the regression of the forecast revision on a long dummy with return controls, year/month fixed effects and controls for multiple hypothesis testing (Column 6 in Table 3).

Table A2: Share of Mispricings among the Predictors - Using only Predictors with significant FDR controlled returns

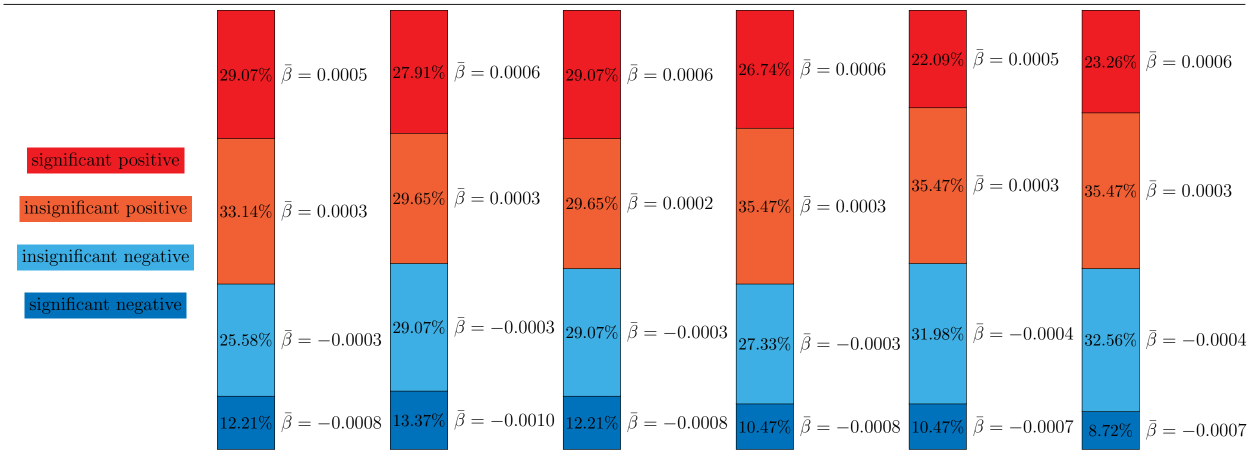
	(1)	(2)	(3)	(4)	(5)	(6)



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	50107	50107	50107	50107	50107	50107
Number of Regressions	105	105	105	105	105	105

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each of the 105 predictors with significant long minus short returns when controlling the false discovery rate to 5%. The table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level. Standard errors are clustered on the firm and the year/month level.

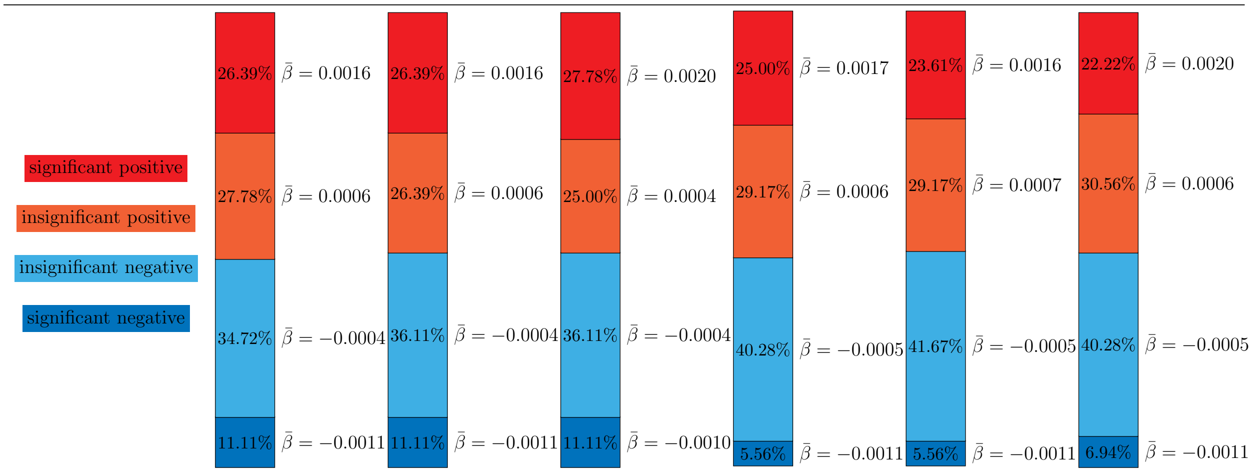
Table A3: Share of Mispricings among the Predictors - Using only 1 Data Point per Portfolio Resorting



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	29586	27024	27024	29586	27024	27024
Number of Regressions	172	172	172	172	172	172

The table shows the result of the regression $forecast\,revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\,revision$. $forecast\,revision_{i,d,t,x}$ is the revision of the consensus five quarter ahead earnings forecast between the time a portfolio is formed and the time of the next resorting. That is, the regression does not use the method to obtain forecast revisions at a monthly frequency described in Section 2.3 but has one data point per portfolio formation. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

Table A4: Share of Build-up & Unresolved Predictors - Errors 2 Years After Portfolio Formation

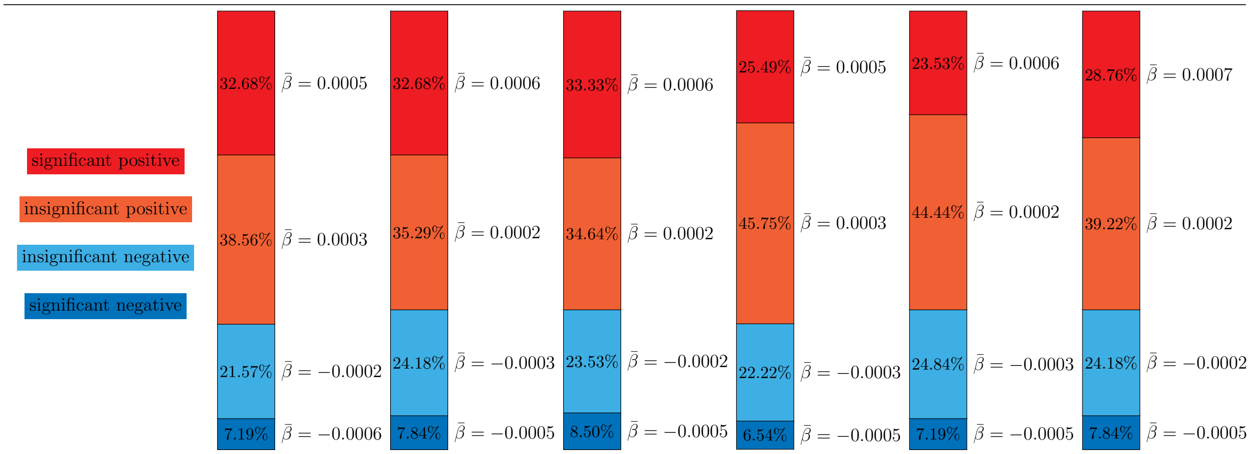


	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	22001	22001	22001	22001	22001	22001
Number of Regressions	72	72	72	72	72	72

The table shows the result of the regression $forecasterror_{i,d,t+24} = \beta_1 + \beta_2 long_{i,t} + X_{i,t} + \epsilon_{i,d,t}$, whereby $forecasterror_{i,d,t+24}$ indicates the forecast error 24 months after portfolio formation. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run for all predictors identified as mispricing in the regression shown in Column 3 of Table 3, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and columns 4-6 control the false discovery rate at 5%. Columns 2,3,5 & 6 control for the return over the period between the formation of the predictor portfolio and the time the forecast was made. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

Table A5: Share of Mispricings among the Predictors - Pre-Publication Data

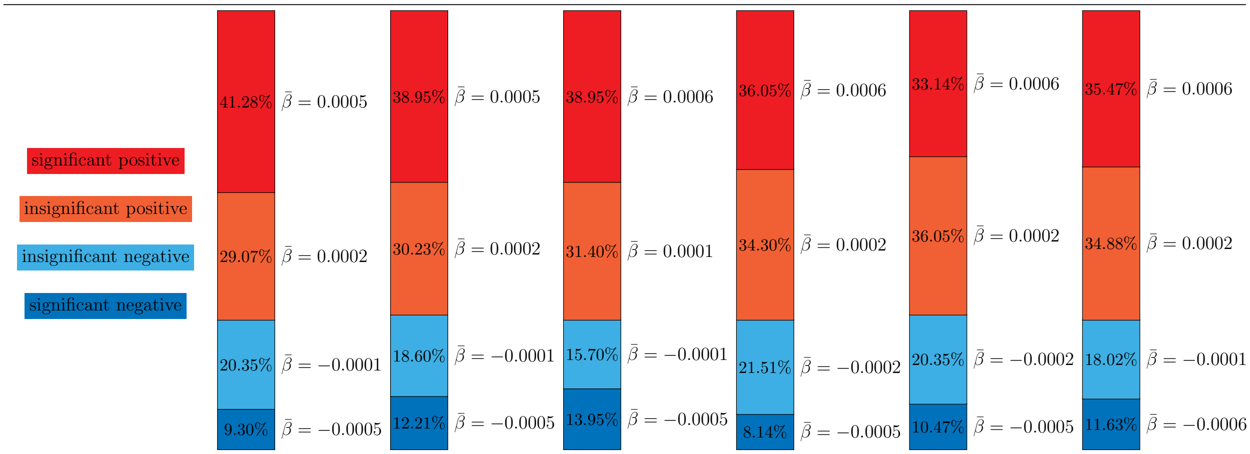
	(1)	(2)	(3)	(4)	(5)	(6)



Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	15703	15703	15703	15703	15703	15703
Number of Regressions	153	153	153	153	153	153

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. The table only shows results for 153 instead of the usual 172 regressions because forecast data only becomes available after 1980, so there is no pre-publication data available for some predictors. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

Table A6: Share of Mispricings among the Predictors - Post-Publication Data



	(1)	(2)	(3)	(4)	(5)	(6)
Control Return	No	Yes	Yes	No	Yes	Yes
Fixed Effects	None	None	Year/Month	None	None	Year/Month
MHT Control	None	None	None	Fdr	Fdr	Fdr
Mean Regression NObs	36600	36600	36600	36600	36600	36600
Number of Regressions	172	172	172	172	172	172

The table shows the result of the following regression: $forecast\ revision_{i,d,t,x} = \beta_1 + \beta_2 long_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,d,t,x}$. $forecast\ revision$. $forecast\ revision_{i,d,t,x}$ is the monthly revision of the consensus five quarter ahead earnings forecast, described in detail in Section 2.3. $long_{i,t}$ is a dummy that is one if a stock is sorted into the long portfolio and zero if it is in the short portfolio. The regression is run separately for each predictor, and the table reports aggregated results. Significance tests are done at the 5% level. Columns 1-3 do not control for multiple hypothesis testing, and Columns 4-6 control the false discovery rate (FDR) at 5%. Columns 2,3,5 & 6 control for the return over the period between the two forecasts included in the revision. Columns 3 & 6 use Year/Month fixed effects. Standard errors are clustered on the firm and the year/month level.

Table A7: Significant Mispricings with Categories

name	t value	parameter	standard error	Description	Economic Category
AnalystRevision	21.22	0.0006	0.000026	EPS forecast revision	earnings forecast
AnnouncementReturn	18.33	0.0006	0.000030	Earnings announcement return	earnings event
ResidualMomentum	16.16	0.0007	0.000043	Momentum based on FF3 residuals	momentum
Mom6mJunk	16.06	0.0011	0.000066	Junk Stock Momentum	momentum
REV6	14.03	0.0007	0.000047	Earnings forecast revisions	earnings forecast
Mom12mOffSeason	13.72	0.0011	0.000084	Momentum without the seasonal part	other
EarningsStreak	13.24	0.0007	0.000051	Earnings surprise streak	earnings growth
DelBreadth	12.86	0.0007	0.000054	Breadth of ownership	ownership
Mom6m	12.08	0.0011	0.000088	Momentum (6 month)	momentum
EarningsForecastDisparity	11.02	0.0005	0.000042	Long-vs-short EPS forecasts	earnings forecast
CompEquIss	9.96	0.0003	0.000031	Composite equity issuance	external financing
NumEarnIncrease	9.96	0.0002	0.000023	Earnings streak length	earnings growth
IdioVol3F	9.40	0.0004	0.000046	Idiosyncratic risk (3 factor)	volatility
RealizedVol	9.34	0.0005	0.000050	Realized (Total) Volatility	volatility
ForecastDispersion	9.27	0.0004	0.000044	EPS Forecast Dispersion	volatility
FEPS	9.10	0.0006	0.000064	Analyst earnings per share	profitability
NetEquityFinance	8.45	0.0012	0.000137	Net equity financing	external financing
RevenueSurprise	8.29	0.0003	0.000034	Revenue Surprise	sales growth
ShareRepurchase	8.11	0.0005	0.000062	Share repurchases	payout indicator
EarningsSurprise	8.02	0.0003	0.000042	Earnings Surprise	earnings growth
MomRev	7.82	0.0008	0.000107	Momentum and LT Reversal	momentum
MomVol	7.77	0.0014	0.000176	Momentum in high volume stocks	momentum
InvestPPEInv	7.49	0.0002	0.000024	change in ppe and inv/assets	investment
dNoa	7.40	0.0002	0.000033	change in net operating assets	investment
Mom12m	7.15	0.0007	0.000103	Momentum (12 month)	momentum
FirmAgeMom	7.03	0.0009	0.000124	Firm Age - Momentum	momentum
XFIN	6.70	0.0012	0.000181	Net external financing	external financing
retConglomerate	6.55	0.0004	0.000059	Conglomerate return	lead lag
roaq	6.22	0.0004	0.000066	Return on assets (qtrly)	profitability
IndRetBig	6.10	0.0004	0.000062	Industry return of big firms	lead lag
ChTax	5.98	0.0003	0.000053	Change in Taxes	other
Tax	5.81	0.0006	0.000108	Taxable income to income	other
DelEqu	5.19	0.0006	0.000106	Change in equity to assets	investment
NetPayoutYield	5.16	0.0013	0.000243	Net Payout Yield	valuation
ShareIss1Y	5.07	0.0005	0.000098	Share issuance (1 year)	external financing
DelCOA	4.69	0.0005	0.000103	Change in current operating assets	investment
ChEQ	4.61	0.0005	0.000106	Growth in book equity	investment
OperProf	4.37	0.0006	0.000139	operating profits / book equity	profitability
IntMom	4.33	0.0003	0.000059	Intermediate Momentum	momentum
CBOperProf	4.26	0.0006	0.000139	Cash-based operating profitability	profitability
zerotradeAlt1	3.88	0.0011	0.000284	Days with zero trades	liquidity
hire	3.80	0.0004	0.000107	Employment growth	investment
IndMom	3.78	0.0002	0.000056	Industry Momentum	momentum
OperProfRD	3.57	0.0006	0.000156	Operating profitability R&D adjusted	profitability
VolumeTrend	3.41	0.0004	0.000123	Volume Trend	volume
AssetGrowth	3.41	0.0006	0.000188	Asset growth	investment
ConvDebt	3.34	0.0001	0.000028	Convertible debt indicator	external financing
CustomerMomentum	3.33	0.0002	0.000065	Customer momentum	lead lag
OScore	3.33	0.0005	0.000158	O Score	default risk
DelFINL	3.29	0.0003	0.000087	Change in financial liabilities	external financing
IndIPO	3.20	0.0001	0.000039	Initial Public Offerings	external financing
RoE	3.18	0.0003	0.000106	net income / book equity	profitability
PctAcc	3.16	0.0005	0.000155	Percent Operating Accruals	accruals
ShortInterest	2.90	0.0002	0.000062	Short Interest	short sale constraints
TotalAccruals	2.84	0.0003	0.000102	Total accruals	investment
PctTotAcc	2.77	0.0004	0.000144	Percent Total Accruals	accruals
Recomm_ShortInterest	2.64	0.0004	0.000155	Analyst Recommendations and Short-Interest	recommendation
EarnSupBig	2.62	0.0002	0.000061	Earnings surprise of big firms	lead lag
InvGrowth	2.51	0.0005	0.000200	Inventory Growth	profitability
MRreversal	2.51	0.0003	0.000113	Medium-run reversal	long term reversal
betaVIX	2.48	0.0001	0.000032	Systematic volatility	volatility
MaxRet	2.47	0.0002	0.000074	Maximum return over month	volatility
iomom_supp	2.46	0.0001	0.000040	Suppliers momentum	lead lag
CoskewACX	2.44	0.0001	0.000043	Coskewness using daily returns	risk
DelNetFin	2.43	0.0002	0.000086	Change in net financial assets	investment
ChInv	2.40	0.0004	0.000150	Inventory Growth	investment
ChForecastAccrual	2.27	0.0002	0.000084	Change in Forecast and Accrual	earnings forecast
CompositeDebtIssuance	2.24	0.0002	0.000088	Composite debt issuance	external financing
NOA	2.21	0.0001	0.000050	Net Operating Assets	asset composition
MomSeason06YrPlus	2.16	0.0001	0.000068	Return seasonality years 6 to 10	other
ChNWC	2.13	0.0002	0.000082	Change in Net Working Capital	investment
EquityDuration	1.99	0.0003	0.000150	Equity Duration	valuation

The table shows the significant mispricings based on the regression of forecast revision on a long dummy with return controls and year/month fixed effects (Column 3 in Table 3), as well as the economic category they belong to based on the classification by A. Y. Chen and Zimmermann (2021).

Table A8: Build-up and Resolution Predictors with Category

name	Economic Category	T Value Start	Significant Fdr Control Start	T Value 1 Year After	Significant Fdr Control 1 Year After
PctAcc	accruals	-0.37	No	-0.13	No
PctTotAcc	accruals	0.77	No	0.64	No
NOA	asset composition	7.19	Yes	3.24	Yes
OScore	default risk	4.68	Yes	2.29	Yes
AnnouncementReturn	earnings event	9.44	Yes	0.49	No
AnalystRevision	earnings forecast	14.44	Yes	-1.38	No
REV6	earnings forecast	5.01	Yes	-7.40	Yes
EarningsForecastDisparity	earnings forecast	-0.29	No	-8.61	Yes
ChForecastAccrual	earnings forecast	0.38	No	0.42	No
EarningsStreak	earnings growth	14.20	Yes	2.67	Yes
NumEarnIncrease	earnings growth	3.30	Yes	-4.05	Yes
EarningsSurprise	earnings growth	4.67	Yes	-4.60	Yes
ShareIss1Y	external financing	1.90	No	3.07	Yes
CompositeDebtIssuance	external financing	0.60	No	0.24	No
ConvDebt	external financing	2.94	Yes	3.26	Yes
XFIN	external financing	2.26	No	5.40	Yes
NetEquityFinance	external financing	4.98	Yes	5.28	Yes
IndIPO	external financing	3.01	Yes	2.66	Yes
DelFINL	external financing	0.90	No	-0.00	No
CompEquIss	external financing	7.66	Yes	1.94	No
TotalAccruals	investment	-1.42	No	-2.76	Yes
hire	investment	0.30	No	-0.55	No
InvestPPEInv	investment	5.74	Yes	1.46	No
DelNetFin	investment	3.12	Yes	1.91	No
ChEQ	investment	0.06	No	-2.14	No
ChInv	investment	-1.58	No	-2.02	No
AssetGrowth	investment	-0.25	No	-1.56	No
DelEqu	investment	-0.13	No	-2.18	No
ChNWC	investment	1.45	No	0.18	No
DelCOA	investment	1.20	No	-1.13	No
dNoa	investment	4.08	Yes	-0.15	No
EarnSupBig	lead lag	0.69	No	-1.97	No
retConglomerate	lead lag	2.87	Yes	-0.64	No
CustomerMomentum	lead lag	4.00	Yes	-0.38	No
IndRetBig	lead lag	1.22	No	-2.70	Yes
iomom_supp	lead lag	2.30	No	-1.78	No
zerotradeAlt1	liquidity	1.59	No	0.68	No
MRreversal	long term reversal	-1.73	No	-0.33	No
MomRev	momentum	-0.88	No	-3.63	Yes
Mom12m	momentum	3.75	Yes	-3.96	Yes
FirmAgeMom	momentum	5.26	Yes	-2.15	No
Mom6m	momentum	4.95	Yes	-3.37	Yes
Mom6mJunk	momentum	9.74	Yes	-5.56	Yes
ResidualMomentum	momentum	12.10	Yes	-2.41	Yes
MomVol	momentum	3.52	Yes	-2.51	Yes
IndMom	momentum	0.51	No	-1.53	No
IntMom	momentum	1.99	No	-4.92	Yes
Tax	other	3.06	Yes	2.21	Yes
ChTax	other	-0.61	No	-4.62	Yes
MomSeason06YrPlus	other	1.65	No	2.32	Yes
Mom12mOffSeason	other	7.32	Yes	-6.81	Yes
DelBreadth	ownership	6.06	Yes	-3.46	Yes
ShareRepurchase	payout indicator	5.65	Yes	8.47	Yes
InvGrowth	profitability	-0.93	No	-0.54	No
OperProfRD	profitability	9.09	Yes	5.45	Yes
RoE	profitability	3.47	Yes	6.98	Yes
FEPS	profitability	7.55	Yes	5.16	Yes
roaq	profitability	4.91	Yes	3.49	Yes
OperProf	profitability	3.61	Yes	5.59	Yes
CBOperProf	profitability	6.37	Yes	4.72	Yes
Recomm_ShortInterest	recommendation	3.76	Yes	3.14	Yes
CoskewACX	risk	0.85	No	-0.53	No
RevenueSurprise	sales growth	0.33	No	-8.11	Yes
ShortInterest	short sale constraints	0.23	No	0.13	No
NetPayoutYield	valuation	3.77	Yes	3.85	Yes
EquityDuration	valuation	1.50	No	2.46	Yes
RealizedVol	volatility	17.16	Yes	10.44	Yes
MaxRet	volatility	11.76	Yes	7.68	Yes
IdioVol3F	volatility	19.38	Yes	12.78	Yes
ForecastDispersion	volatility	16.63	Yes	10.13	Yes
betaVIX	volatility	0.62	No	-1.62	No
VolumeTrend	volume	0.65	No	0.56	No

The table shows the results of a regression of the forecast errors of five quarters ahead forecasts made at the time of portfolio formation (Column 3 in Table 8) and the results of a regression of the forecast error of five quarters ahead forecasts made one year after portfolio formation (Column 3 in Table 9) on a long dummy, for each predictor that is classified as a mispricing based on the regression of forecast revision on a long dummy (Column 3 in Table 3), as well as the economic category they belong to based on the classification by A. Y. Chen and Zimmermann (2021).

Table A9: Significant Mispricings - Comparison to Holcblat et al. (2022)

Name	P-value Expectations	P-value Risk
AnalystRevision	p<0.001	p=1.000
AnnouncementReturn	p<0.001	p=1.000
AssetGrowth	p<0.001	p=1.000
betaVIX	p=0.013	p=0.400
CBOperProf	p<0.001	p=1.000
ChEQ	p<0.001	p=0.640
ChForecastAccrual	p=0.023	p=0.400
ChInv	p=0.016	p=1.000
ChTax	p<0.001	p=0.580
CompEquIss	p<0.001	p=1.000
CompositeDebtIssuance	p=0.025	p=1.000
ConvDebt	p<0.001	p=1.000
CoskewACX	p=0.015	p=0.360
CustomerMomentum	p<0.001	p=0.270
DelBreadth	p<0.001	p=0.590
DelCOA	p<0.001	p=1.000
DelEqu	p<0.001	p=0.730
DelFINL	p<0.001	p=1.000
DelNetFin	p=0.015	p=1.000
dNoa	p<0.001	p=1.000
EarningsForecastDisparity	p<0.001	p=0.370
EarningsStreak	p<0.001	p=1.000
EarningsSurprise	p<0.001	p=0.470
EarnSupBig	p=0.009	p=0.530
FEPS	p<0.001	p=1.000
FirmAgeMom	p<0.001	p=1.000
ForecastDispersion	p<0.001	p=1.000
hire	p<0.001	p=1.000
IdioVol3F	p<0.001	p=1.000
IndIPO	p=0.001	p=0.470
IndMom	p<0.001	p=0.250
IndRetBig	p<0.001	p=1.000
IntMom	p<0.001	p=0.490
InvestPPEInv	p<0.001	p=0.050
InvGrowth	p=0.012	p=1.000
iomomsupp	p=0.014	p=1.000
MaxRet	p=0.014	p=0.420
Mom12m	p<0.001	p=1.000
Mom12mOffSeason	p<0.001	p=1.000
Mom6m	p<0.001	p=1.000
Mom6mJunk	p<0.001	p=0.530
MomRev	p<0.001	p=1.000
MomVol	p<0.001	p=1.000
NetEquityFinance	p<0.001	p=1.000
NetPayoutYield	p<0.001	p=1.000
OScore	p<0.001	p=1.000
PctAcc	p=0.002	p=0.460
PctTotAcc	p=0.006	p=0.490
RealizedVol	p<0.001	missing
RecommShortInterest	p=0.008	missing
ResidualMomentum	p<0.001	p=1.000
retConglomerate	p<0.001	p=1.000
REV6	p<0.001	p=1.000
RevenueSurprise	p<0.001	p=0.300
roaq	p<0.001	p=1.000
ShareIss1Y	p<0.001	p=1.000
Tax	p<0.001	p=0.370
TotalAccruals	p=0.005	p=0.030
XFIN	p<0.001	p=1.000
zerotradeAlt1	p<0.001	p=1.000
MRreversal	p=0.012	p=1.000
NumEarnIncrease	p<0.001	p=1.000
OperProf	p<0.001	p=1.000
OperProfRD	p<0.001	p=1.000
RoE	p=0.001	p=1.000
ShareRepurchase	p<0.001	p=1.000
ShortInterest	p=0.004	p=1.000
VolumeTrend	p<0.001	p=1.000

The table compares the expectation-based mispricing specification from this paper and the risk-based classification by Holcblat et al. (2022). The Column “P-value Expectations” shows the results of the regression of forecast revision on a long dummy with return controls and year/month fixed effects (Column 3 in Table 3). A low p-value indicates that the predictor is a mispricing. The column “P-value Risk” shows the result of their unconditional test (Table A.11 of the march 2022 version of their working paper). A high p-value indicates that the predictor is a mispricing.

E Detailed Predictor Definitions

This section contains more detailed variable definitions for the predictors mentioned by name in the main part of the paper. Definitions come from the Open Source Asset Pricing (OpenAP) data documentation by A. Y. Chen and Zimmermann (2021) and are in some cases augmented with details from the original paper or highlight differences between the OpenAP definition and the one in the original paper. The OpenAp definitions are in quotation marks. Anything not in quotation marks is additional information.

AbnormalAccruals "Define Accruals as net income (ib) minus operating cash flow (oancf), divided by average total assets (at) for years t-1 and t. If oancf is missing, replace operating cash flow with funds from operations (fopt) minus the annual change in total current assets (act) plus the annual change in cash and short-term investments (che) plus the annual change in current liabilities (lct) minus the annual change in debt in current liabilities (dlc). For each year t and 2-digit sic code, regress Accruals on: the inverse of average total assets for year t-1, the change in revenue (sale) from year t-1 to t divided by total assets for t-1, property plant and equipment (ppegt) divided by total assets for t-1. AbnormalAccrual is the residual from this cross-sectional regression." Originally from Xie (2001).

Accruals "Annual change in current total assets (act) minus annual change in cash and short-term investments (che) minus annual change in current liabilities (lct) minus annual change in debt in current liabilities (dlc) minus change in income taxes (txp). All divided by average total assets (at) over this year and last year. Exclude if abs(prc) < 5." Originally by Sloan (1996), who interprets this predictor as resolution of mispricing.

AccrualsBM - Book-to-market and accruals "Binary variable equal to 1 if stock is in the highest Accrual quintile and the lowest BM quintile, and equal to 0 if stock is in the lowest Accrual quintile and the highest BM quintile. Exclude if book equity (ceq) is negative." Originally from Bartov and Kim (2004).

AdExp "Advertising expense (xad) over market value of equity (shrout*abs(prc))" Originally from Chan et al. (2001).

AnalystRevision - EPS forecast revision "keep fpi == 1, last obs each month. Signal is meanest / last month's meanest." Take stocks that are followed by three or more analysts that have shown the greatest increase in their mean earnings forecasts since the prior month. The horizon of the forecasts is not specified in the paper. Original paper by Hawkins et al. (1984).

AnnouncementReturn - Earnings announcement return "Get announcement date for quarterly earnings from IBES (fpi = 6). AnnouncementReturn is the sum of (ret - mktrf + rf) from one day before an earnings announcement to 2 days after the announcement." Stocks with positive AnnouncementReturn are in the long portfolio. Originally by Chan et al. (1996).

AssetGrowth "Annual growth rate of total assets (at)". Firms that grow less are in the long portfolio. Originally by Cooper et al. (2008).

BetaLiquidityPS "Monthly excess return (ret -rf) regressed on innovations in liquidity from Pastor's website (`https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt`). Use 60 month rolling window regression, and require at least 36 non-missing observations." Originally from Pástor and Stambaugh (2003)

BMdec - Book to market using December Market Equity "BM using most recent December value of market equity." Originally from Fama and French (1992).

Hire - Employment growth "Change in number of employees (emp) between t -1 and t, scaled by average number of employees in t-1 and t. Replace hire with 0 if emp or lagged emp is missing. Firms with fewer hires are in the long portfolio." The original paper by Belo et al. (2014) has a model and offers a risk-based explanation for this effect.

ChEQ - Growth in book equity "Ratio of book equity (ceq) to book equity in the previous year. Include only if book equity is positive this year and last year." Originally by Lockwood and Prombutr (2010).

ChInv - Inventory Growth "12 month change in inventory (invt) divided by average total assets." Originally by Thomas and H. Zhang (2002). Change in inventory is a component of Accruals.

ChInvIA - Change in capital inv (ind adj) "Growth in capital expenditure (capx) minus average growth in capital expenditure in the same industry (two-digit SIC). If capx is missing, capital expenditure is defined as the annual change in property, plant and equipment (ppent). Capital expenditure growth is defined as the percentage growth of capx today relative to the average capx over the previous two years (.5*(capx_{t-1} + capx_{t-2}), or as percentage growth relative to the previous year only if t-2 is missing." Stocks with low growth in capital expenditure are in the long portfolio. Originally by Abarbanell and Bushee (1998), who use a regression and initially start with the opposite hypothesis that more investment should mean higher returns.

ChTax - Change in Taxes "4-quarter change in quarterly total taxes (txtq), scaled by lagged total assets (at)." Stocks with an increase in tax expense are in the long portfolio. Originally by Thomas and H. Zhang (2002), whose interpretation is that an increase in tax, while it may seem bad, is a proxy for increased core profitability to which the market underreacts.

CompEquIss - Composite Equity Issuance "5 year growth rate of market value of equity minus 5 year stock return." The long portfolio contains the stocks with low Composite Equity Issuance. Originally by Daniel and Titman (2006).

DelCOA - Change in current operating assets "Difference in current operating assets (total current assets (act) minus cash and short-term investments (che)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005). The long portfolios are the stocks with a low difference in operating assets.

DelBreadth - Breadth of ownership "Quarterly change in the number of institutional owners (numinstowners) from 13F data. Exclude if in the lowest quintile of stocks by market value of equity (based on NYSE stocks only)." Stocks with positive changes in owners are in the long portfolio. Originally from J. Chen et al. (2002).

DelEqu - Change in equity to assets "Difference in book equity (ceq) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005).

DelFINL - Change in financial liabilities "Difference in financial liabilities (sum of long-term debt (dltt), current liabilities (dlc) and preferred stock (pstk)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t." Originally by Richardson et al. (2005). They do Fama-MacBeth regression and find a significant positive effect on return if only the variable is included and a negative (insignificant) if it is included together with their other financing measures. The original paper defines the measure in the opposite way as the difference between years t and t-1, but then flips the sign for the relevant regression in table 8. Stocks with a decrease in liabilities are in the long portfolio.

dNoa - change in net operating assets "12-month growth in Net Operating Assets scaled by lagged total assets (at). Net Operating assets are operating assets minus operating liabilities. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dlc), book equity (ceq) and preferred stock (pstk), all items (except at and ceq) replaced with 0 if missing." Stocks with a decrease in NOA are in the long portfolio. Originally by Hirshleifer et al. (2004). They find a strong effect in a regression but only when not controlling for the level of net operating assets, which is also among the predictors (NOA)

EarningsForecastDisparity - Long-vs-short EPS forecasts "Analyst forecasted 5-year earnings growth (fgr5yr) minus 100 times the difference between mean earnings forecast (meanest) and fiscal year earnings expectations (fy0a). " The original paper (Da and Warachka, 2011) uses the following definition, which is somewhat different: Compare 1-year growth forecast ISTG calculated by taking IBES 1 year ahead forecast and getting the change to the last announced value to LTG. The short portfolio is the one that has a high LTG and low ISTG, and the long portfolio has a Low LTG and high ISTG.

EarningsStreak - Earnings surprise streak "Use fpi == 6 and only the last statpers for each anndats_act. Define surp = (actual - meanest)/price. Define a firm-anndats as a streak if surp has the same sign as the most recent surp observation. Keep only streaks. Then define signal = surp." Originally by Loh and Warachka (2012).

FEPS - Analyst Earnings Per Share "Using IBES unadjusted forecasts, keep fpi == 1, signal is meanest." Construct the mean 1 Year ahead earnings per share forecast in the previous month. The original paper (Cen et al., 2006) first sorts on firm characteristics (size, price BTM and 6-month return) and then within these sorts on FEPS. Stocks with high forecasts are in the long portfolio.

FirmAgeMom - Firm Age - Momentum "6 month return, restricted to the bottom quintile of the cross-sectional firm age distribution. Exclude if price less than 5 or firm younger than 12 months." Originally from X. F. Zhang (2006).

ForecastDispersion - EPS Forecast Dispersion "Keep fpi = 1 and fpedats ¿ statpers + 30. Standard deviation of earnings estimates (stdev_est) scaled by mean earnings estimate." "Dispersion is defined as the standard deviation of earnings forecasts scaled by the absolute value of the mean earnings forecast. If the mean earnings forecast is zero, then the stock is assigned to the highest dispersion category." Use one-year-ahead forecasts. Originally from Diether et al. (2002)

Frontier - Efficient frontier index "Frontier is the residual of a regression of log(BM) on log(book equity (ceq)), long-term debt (dltt) to assets (at), capital expenditures (capx) to revenue (sale), R&D expense (xrd) to revenue, advertising expense (xad) to revenue, property plant and equipment (ppent) to assets, EBIT (ebitda) to assets, and dummies for Fama-French's 48 industry definitions. Regression is updated each month with a rolling window of 60 months." Originally from Nguyen and Swanson (2009).

grcapx - Change in capex (two years) "Growth rate of capital expenditures (capx) relative to two years ago. If capx is missing, replace with annual change in property, plant and equipment (ppent)." Originally from Anderson and Garcia-Feijóo (2006).

grcapx3y - Change in capex (three years) "Capital expenditures (capx) divided by the sum of capital expenditures from year - 1, year -2, and year -3. If capx is missing, replace with annual change in property, plant and equipment (ppent)." Originally by Anderson and Garcia-Feijóo (2006).

Illiquidity "Past twelve month average of: daily return (abs(ret)) divided by turnover((abs(prc)*vol)" Originally from Amihud (2002).

IndRetBig - Industry return of big firms "Average monthly return (ret) of the 30% largest companies by market value of equity in the same Fama-French 48 industry. Exclude the largest 30% of companies for IndRetBig (not to compute the anomaly!) " Originally from Hou (2007)

IntMom - Intermediate Momentum "Stock return between months t-12 and t-6" Originally from Novy-Marx (2012).

IO ShortInterest - Institutional ownership among high short interest stock "Exclude all stocks with short interest (ShortInterest) below 99th percentile. IO_ShortInterest is institutional ownership (instown_perc). Keep NYSE Only." Originally from Asquith et al. (2005).

InvestPPEInv - change in ppe and inv/assets "One-year change in property, plants and equipment (ppegt) plus one year change in inventory (invt), scaled by one-year lagged assets (at)." Originally by Lyandres et al. (2008).

Leverage - Market leverage "Total liabilities (lt) divided by market value of equity." Stocks with high leverage are in the long portfolio. Originally by Bhandari (1988)

MeanRankRevGrowth - Revenue Growth Rank "Rank firms by their annual revenue growth each year over the past 5 years. MeanRankRevGrowth is the weighted average of ranks over the past 5 years, that is, MeanRankRevGrowth = (5*Rank_{t-1} + 4*Rank_{t-2} + 3*Rank_{t-3} + 2*Rank_{t-4} + 1*Rank_{t-5})/15. Exclude NASDAQ stocks." Stocks with low growth are in the long portfolio. Originally from Lakonishok et al. (1994), who use a somewhat different methodology that sorts on multiple characteristics.

Mom6m - Momentum (6 month) "Stock return between months t-6 and t-1" following Jegadeesh and Titman (1993).

Mom6mJunk - Junk Stock Momentum "Mom6m. Include only stocks with a credit rating (splticrm) of BBB or lower" Originally by Avramov et al. (2007).

Mom12m - Momentum (12 month) "Stock return between months t-12 and t-1" Originally from Jegadeesh and Titman (1993).

MomOffSeason - Off-season long-term reversal "Average return in other months over the preceding 2-5 years." It is the counterpart to "MomSeason". Originally from Heston and Sadka (2008).

Mom12mOffSeason - Momentum without the seasonal part "Average return in other months over the previous year." Originally from Heston and Sadka (2008).

MomSeason - Return seasonality years 2 to 5 "Average return in the same month over the preceding 2-5 years." Originally by Heston and Sadka (2008).

MomVol - Momentum in high volume stocks "Define momentum as Mom6m, and volume as the rolling average of the past 6 months of monthly turnover (minimum 5 months). Independent sort stocks into 10 momentum ports and 3 volume ports. Keep if volume is in the top port, and assign signal = momentum port. Drop if less than 2 years on CRSP." Originally by Lee and Swaminathan (2000).

MRreversal - Medium-run reversal "Stock return between months t-18 and t-13." Based on De Bondt and Thaler (1985) who do not do a statistical test for this horizon and just provide visual evidence.

NetEquityFinance "Sale of common stock (sstk) minus purchase of common stock (prstkc), scaled by average total assets (at) from years t and t-1. Exclude if absolute value of ratio is greater than 1." Stocks with high net equity financing are in the short portfolio, and those with low net equity finance are in the long portfolio. Originally by Bradshaw et al. (2006).

NetPayoutYield "Dividends (dvc) plus purchase of common and preferred stock (prstkc) minus sale of common and preferred stock (sstk), divided by market value of equity lagged 6 months. Exclude if NetPayoutYield is 0, financial firm based on SIC code, ceq ¡= 0, or less than 2 years in CRSP". High-yield stocks are in the long portfolio, and low-yield stocks are in the short portfolio. Originally by Boudoukh et al. (2007).

NOA - Net Operating Assets "Difference between operating assets and operating liabilities, scaled by lagged total assets. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dc) and book equity (ceq)." Stocks with low NOA are in the long portfolio. Originally by Hirshleifer et al. (2004).

NumEarnIncrease - Earnings streak length "Number if consecutive 4-quarter increases in ibq, up to 8." Stocks with a long positive streak are in the long portfolio. This is based on Loh and Warachka (2012). They do not study the same signal, but their Table 4 suggests this should work. Their definition of a streak is just that the surprise has the same sign (but it does not have to increase). Their Table 4 shows that earnings surprises have a stronger effect if they are part of a streak rather than a reversal of a streak.

PctTotAcc - Percent Total Accruals "Net income (ni) minus (purchase of common and preferred stock (prstkcc) minus sale of common and preferred stock (sstk) plus dividends (dvt), cash flow from operations (oancf), from financing (fincf) and investment (ivncf)). Scaled by absolute value of net income." Stocks with low values of the characteristic are in the long portfolio. Originally by Hafzalla et al. (2011).

PriceDelayRsq - Price delay r square "Each July regress stock return on day t on market return in $t, t-1, \ldots, t-4$ using observations from July 1 of the previous year to June 30 of the current year. Then regress again with no market return lags. PriceDelay Rsq = 1 - [Rsq from second regression]/[Rsq from first regression]" Firms with higher delay are in the long portfolio. Originally from Hou and Moskowitz (2005). Measures delayed market response to information, as it captures how much of the market level information is incorporated into the stock price on the same day compared to on later days. If everything is incorporated on the same day, adding the lagged returns does not improve R^2 subsequentially.

RealizedVol - Realized (Total) Volatility "Standard deviation of residuals from CAPM regressions using the past month of daily data. Value weighted" Originally by Ang et al. (2006).

Recomm_ShortInterest - Analyst Recommendations and Short-Interest "Go long firms in lowest quintile of short interest (shortint/shrout) and lowest quintile of analyst recommendations (monthly consensus recommendation using the most recent analyst recommendation in the past 12 months). Go short firms in highest quintile of short interest and highest quintile of analyst recommendations." Originally by Drake et al. (2011). Note that the construction of the portfolios is different from the original paper due to a lack of data.

retConglomerate - Conglomerate return "Identify conglomerate firms as those with multiple OPSEG or BUSSEG entries in the Compustat segment data (and require that at least 80% of firm's total assets are covered by segment data). Compute monthly stock return at the 2-digit SIC level for stand-alone (non-conglomerate) firms only, and match those returns to conglomerates' segments. Compute weighted conglomerate return as the industry return of stand-alone companies, weighted with a conglomerate's total sales in each industry." Originally by Cohen and Lou (2012).

RDIPO - IPO and no R&D spending "Binary variable equal to 1 if R&D expense (xrd) = 0 and IndIPO = 1. 0 otherwise." Stocks with RnD are in the long portfolio, and stocks without are in the short portfolio. Originally from R.-J. Guo et al. (2006).

RDS - Real Dirty Surplus "Define Dirty Surplus as annual change in marketable securities adjustment msa plus annual change in retained earnings adjustment (recta) + .65 times the annual change in min(Unrecognized prior service cost (pcupsu) - Pension additional minimum liability (paddml),0). Real dirty surplus is the annual change in book equity (ceq) minus dirty surplus minus (net income (ni) minus dividends preferred (dvp)) + dividends (divamt) - end-of-fiscal-year-stock-price (prcc_f)*annual change in common shares outstanding (csho)." Originally from Landsman et al. (2011).

roaq - Return on assets (qtrly) "Quarterly net income (ibq) divided by lagged total assets (atq)." Originally by Balakrishnan et al. (2010).

ShareIss1Y - Share issuance (1 year) "Growth in number of shares between t-18 and t-6. Number of shares is calculated as shrout/cfacshr to adjust for splits." Stocks with the most positive share issuance are in the short portfolio and those with the most negative share issuances are in the long portfolio. The original paper by Pontiff and Woodgate (2008) uses Fama-MacBeth regressions instead of portfolio sorts.

ShareRepurchase "Binary variable equal to 1 if stock repurchase indicated in cash flow statement (prstkc > 0), and 0 if prstkc = 0." Stocks with share repurchases form the long portfolio, and stocks without form the short portfolio. The intuition from the original paper by Ikenberry et al. (1995) is that firms strategically buy back shares if they are undervalued. The original paper hand collects share repurchase data by looking at all share repurchases announced in the *Wall Street Journal*.

sinAlgo - Sin Stock A dummy that is one if at least one segment of a firm belogs to an industry based on the NAICS code that is a so-called "sin industry" and zero otherwise. Originally from Hong and Kacperczyk (2009).

Size "Log of monthly market value of equity (abs(prc)*shrout))." Originally from Banz (1981).

SmileSlope - Put volatility minus call volatility "Using OptionMetrics's daily volatility surfaces (vsurfd), keep last observation each month, delta = 0.50 or -0.50, and days to expiration = 30. The signal is then the difference between put implied vol and call implied vol." Originally from Yan (2011).

SP - Sales-to-price "Ratio of annual sales (sale) to market value of equity." Originally from Barbee et al. (1996).

STreversal - Short-term reversal "Stock return (ret) over the previous month." Originally from Jegadeesh (1990)

Tax "Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5." Originally from Lev and Nissim (2004).

TotalAccruals "Before 1988: Change in net working capital ((act - che) - (lct - dlc)) plus change in net noncurrent assets ((at - act - ivao) - (lt - dlc - dltt)) plus change in net financial assets ((ivst + ivao - (dltt + dlc + pstk)). Starting in 1988: net income (ni) minus total, operating and investment cashflows (oancf, ivncf, fincf) plus stock sales minus repurchases and dividends (sstk, prstkc, dv)). Scaled by lagged total assets (at). Replace missings in ivao, ivst, dltt, dlc, pstk sstk, prstkc, dv with 0." Originally from Richardson et al. (2005).

Trend Factor "Normalised moving average of the stockprice" Originally from Han et al. (2016)

VolSD "Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5." Originally from Lev and Nissim (2004)

XFIN - Net external financing "Sale of common stock (sstk) minus dividends (dv) minus purchase of common stock (prstkc) plus long-term debt issuance (dltis) minus long-term debt reductions (dltr). Scaled by total assets (at)." Stocks with high net external financing are in the short portfolio and those with low net external finance are in the long portfolio. Originally form Bradshaw et al. (2006).

zerotrade - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48{*}10^5$), multiplied by 21/number of trading days per month. Zerotrade is the 6-month average of that variable." Originally from W. Liu (2006).

zerotradeAlt1 - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48{*}10^5$), multiplied by 21/number of trading days per month. Take 1-month average." Originally from W. Liu (2006).

zerotradeAlt12 - Days with zero trade "In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48{*}10^5$), multiplied by 21/number of trading days per month. Take 12-month average." Originally from W. Liu (2006).