Expected Mispricing*

Turan G. Bali† Heiner Beckmeyer‡ Timo Wiedemann§

This version: September 6, 2024

Abstract

Following the idea that arbitrageurs care as much about the timing of when an identified mispricing opportunity resolves as they do about the level of mispricing, we introduce a novel return-based measure of firm-level expected mispricing. We define mispricing as the deviation from the fair rate-of-return implied by a statistically optimized factor model and construct our out-of-sample measure of *expected* mispricing, $\mathbb{E}[MP]$, using machine learning models. $\mathbb{E}[MP]$ outperforms existing approaches in predicting future realized mispricing and produces a monthly FF6-alpha of 1.5% in value-weighted portfolios. We show that the degree of mispricing is closely related to limits-to-arbitrage and informational frictions.

Keywords: Mispricing, Machine Learning, IPCA, Limits to Arbitrage

JEL classification: G12, G14, G17, C1

*We thank Bryan Kelly and Nicole Branger for insightful comments and discussions. We are also grateful to Jianan Liu for providing SAS code to calculate the Stambaugh, Yu, and Yuan (2015) mispricing scores and Yi Tang for code on the analyst earnings forecast dispersion measure of Diether, Malloy, and Scherbina (2002).

†Robert S. Parker Chair Professor of Finance, McDonough School of Business, Georgetown University, Washington, D.C. 20057. Phone: +1 (202) 687-5388, Email: turan.bali@georgetown.edu.

‡University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany. Phone: +49 (251) 83-22771. Email: heiner.beckmeyer@wiwi.uni-muenster.de.

§University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany. Phone: +49 (251) 83-22885. Email: timo.wiedemann@wiwi.uni-muenster.de.

1. Introduction

An accurate assessment of mispricing is vital to an efficient allocation of capital. If arbitrageurs can distinguish between overvalued and undervalued firms, they can deploy their funds accordingly, which resolves the mispricing and grants the arbitrageur a profit for putting her capital at risk. Market prices become more efficient and provide a more accurate characterization of a firm's future prospects. The literature has proposed a multitude of mispricing proxies, each with its own benefits and drawbacks. We introduce an overarching mispricing measure, which is not only grounded in economic theory, but also satisfies the requirements of arbitrageurs for a timely and accurate resolution of mispricing.

We contribute to the literature by proposing a novel measure of *expected mispricing*, which is theoretically motivated and consistent with arbitrageurs' trading activity. We define mispricing as the deviation from the fair price implied by the stochastic discount factor. Since the true stochastic discount factor is inherently unobservable, we approximate it with a statistically optimized factor model derived from the return data itself. Kelly, Pruitt, and Su (2019) show that *instrumented principal component analysis* (IPCA) significantly outperforms common observable factor models in describing the risk-return trade-off in the cross-section of individual stocks. Furthermore, the model specification alleviates the need for researchers to pre-specify which factors will work for a given time period (McLean and Pontiff, 2016). We validate that IPCA uncovers common sources of return variation better than competing factor models, and hence compute a stock's *realized mispricing* based on the IPCA-based expected return benchmark. Based on the intuition that mispricing should resolve over time, we characterize a stock as overvalued whenever its realized excess return is too small relative the fair (expected) rate of return implied by the factor model. In a second step, we use a battery of modern machine

learning estimators to predict the future realized mispricing with a large set of firm characteristics as inputs. We employ neural networks with three hidden layers, which Gu, Kelly, and Xiu (2020) identify as the best machine learning specification to predict future stock returns, as well as random forests (Breiman, 2001) and gradient-boosted regression trees (Friedman, 2001). We form an equal-weighted ensemble, which balances out the individual benefits and drawbacks of the three prediction methods, and yields a stabler fit.[1] The output of the machine learning ensemble is thus the *expected mispricing* measure $\mathbb{E}[\text{MP}]$. We compute $\mathbb{E}[\text{MP}]$ based on a rolling estimation window, carefully assuring that a hypothetical investor seeking to identify mispriced investment opportunities in the market can do so in real-time.

Our definition of mispricing closely follows the real-time investment execution process of arbitrageurs. Arbitrage capital is typically raised by hedge funds, focusing on two things when it comes to benefiting from mispricing: first, that the resolution of mispricing can be adequately timed. While price-based mispricing measures, for example, those proposed by Bartram and Grinblatt (2018) or van Binsbergen, Boons, Opp, and Tamoni (2023), have been shown to predict future returns, they are silent about the timing of the resolution of mispricing.[2] Accordingly, van Binsbergen et al. (2023) write that "the real economic impact of an anomaly depends [...] on the timing of its resolution" and present evidence that price wedges persist even after five years. Consequently, this can bind arbitrage capital for extended periods of time, exposing the arbitrageur to considerable timing risk. John M. Keynes already noted in the 1930s that "[m]arkets can stay irrational longer than you can stay solvent." Our mispricing proxy measures the expected resolution of mispricing over the next month – a natural time frame for institutional investors. Second,

[1] Rapach, Strauss, and Zhou (2010) provide an early example of the benefits of ensembles in finance.

[2] The study by Bartram and Grinblatt (2018) shows that price deviations implied by fundamental analysis can predict future stock returns. van Binsbergen et al. (2023) use a price wedge measure to classify return anomalies as those that exacerbate the mispricing of many stocks vs. those that resolve it. Our focus is to study the economics of stock-level mispricing. In Section 4.5 we present empirical evidence consistent with the findings of van Binsbergen et al. (2023).

hedge fund managers are interested in limiting their exposure to common sources of return variation. Our definition of (expected) mispricing is orthogonal to a set of latent risk factors, which Kelly et al. (2019) show to describe the risk-return trade-off among stocks substantially better than leading observable factor models. Of course, the latent factors potentially capture systematic mispricing, a circumstance shared with observable factor models (Kozak, Nagel, and Santosh, 2018).[3] All the same, our definition of mispricing cleanly delineates idiosyncratic mispricing of a target stock from the stock's exposure to systematic variation. Other prominent mispricing proxies, for example, those proposed by Stambaugh et al. (2015) and Engelberg, McLean, and Pontiff (2018), invariably capture systematic (risk) exposure. This in turn exposes the arbitrageur to systematic risk, which greatly elevates the non-hedgeable risk profile of her portfolio.

In Figure 1, we provide insights into the behavior of our mispricing measure by showing its temporal evolution for two stocks, which have received considerable attention in the (financial) press and on social media platforms like Reddit or SeekingAlpha: GameStop Corp. (ticker GME) and Tesla, Inc. (ticker TSLA). We compare the resulting expected mispricing ($\mathbb{E}[\text{MP}]$) shown on the left axes with the Stambaugh et al. (2015) mispricing score (SYY) shown on the right. For ease of comparison, we show the value of $100-\text{SYY}$, such that a larger reading for both measures indicates an expected undervaluation of the stock. This circumstance already highlights a further benefit of $\mathbb{E}[\text{MP}]$ as compared to SYY: since we express $\mathbb{E}[\text{MP}]$ as the expected mispricing return component of a stock, we can make a direct statement about the absolute degree of over- vs. undervaluation of a firm. Thus, a positive (negative) $\mathbb{E}[\text{MP}]$ directly shows that a stock is expected to be undervalued (overvalued). SYY, on the other hand, is expressed relative to the prevailing cross-section, so it is hard to tell, if for example TSLA is undervalued in an absolute sense or "only" relative to other stocks in the market.

[3] Kozak et al. (2018) show that it is impossible to fully discern risk-based from mispricing-based explanations for reduced-form risk factors, which are commonly used in empirical asset pricing.

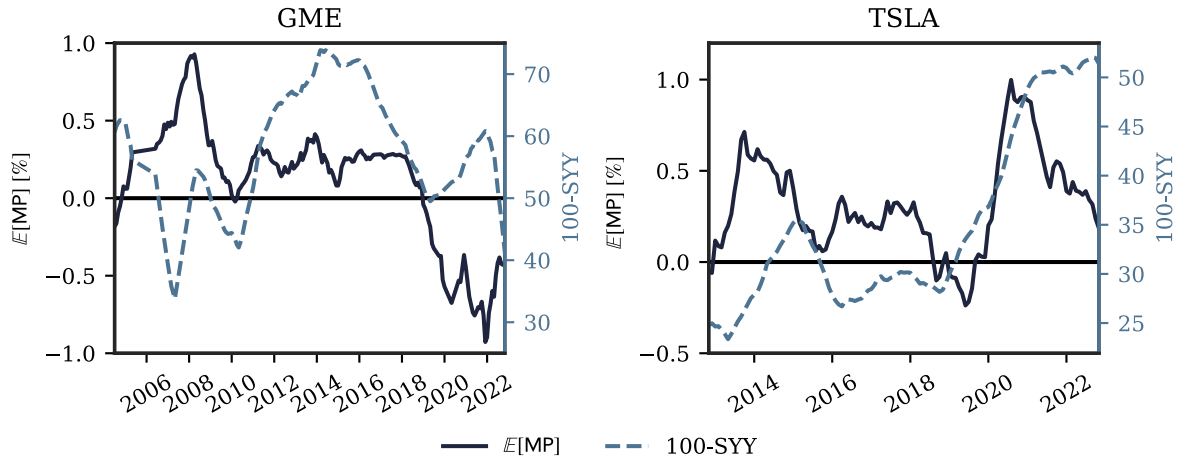


Fig. 1. Time-series of Expected Mispricing for GME and TSLA

The figure shows rolling 12-month averages of expected mispricing for both Gamestop (GME) and Tesla Inc. (TSLA). We compare our measure of expected mispricing ($\mathbb{E}$[MP]) to the Stambaugh et al. (2015) mispricing score (SYY). For ease of comparison, SYY is expressed as 100 − SYY, such that a larger reading indicates an expected undervaluation for both measures.

The left panel in Figure 1 shows that $\mathbb{E}$[MP] considers GME undervalued before and during the financial crisis with a reversion to GME's fair valuation between 2010 and 2018. After that, we find that GME is already viewed as overvalued and interestingly that a large portion of its expected overvaluation is *corrected* after the initial hype around GameStop and the coordinated short squeeze attempt of Redditors. This finding is consistent with a recent study by Zhou and Zhou (2023), who show that institutional investors were for the most part responsible for the dramatic surge in GameStop's stock price. A notable point in Figure 1 is that for a large part of the sample, SYY tends to disagree with $\mathbb{E}$[MP]. Starting in 2005, the two mispricing proxies demonstrate contrasting movements and only begin to co-move after 2012. This remains until around 2019, where we observe that $\mathbb{E}$[MP] commences to pick up on the perceived overvaluation of GME. SYY, on the contrary, only begins to reveal a potential overvaluation of GME in 2022.

The right panel in Figure 1 shows that TSLA is expected to be undervalued until 2019 and afterwards enters a brief period of relative overvaluation. In the second half of 2019 and especially with the Covid-induced selloff of March 2020, we find that $\mathbb{E}$[MP] considers

TSLA once more to be undervalued. This undervaluation is corrected over the following years. The dynamics of SYY tend to agree with the valuation of $\mathbb{E}$[MP] and look like a smoothed and slightly delayed version of it. However, SYY does not show the resolution of TSLA's undervaluation.

We show that our setup identifies mispricing for the overall U.S. stock market considerably well, over the 30-year long out-of-sample testing period between 1993 and 2022. A decile spread portfolio – constructed with our $\mathbb{E}$[MP] measure – that takes a long position in the most undervalued stocks and a short position in the most overvalued stocks earns a value-weighted realized mispricing return of 2.13% per month with an annualized information ratio of 1.6. This long-short portfolio of stocks sorted by $\mathbb{E}$[MP] also produces a monthly value-weighted alpha of 1.5% with respect to the six-factor model of Fama and French (2018) with the market, size, value, investment, profitability, and momentum factors. Consistent with the intuition that mispricing is most rampant among smaller stocks, we find that the realized mispricing, the information ratio, and the alpha spread increase substantially as we move from value-weighted to equal-weighted portfolios. Overvalued stocks tend to be smaller, illiquid and less profitable, and attract more uninformed demand through a lower nominal price (Birru and Wang, 2016) and higher MAX return (Bali, Cakici, and Whitelaw, 2011). Furthermore, analysts disagree significantly more about the future earnings of overvalued stocks (Miller, 1977). Both over- and undervalued stocks have a higher idiosyncratic volatility, agreeing with the argument by Stambaugh et al. (2015) that idiosyncratic volatility proxies for arbitrage costs: stocks with costlier arbitrage activity are more likely to be mispriced according to $\mathbb{E}$[MP].

As our second contribution, we investigate the economics of stock-level mispricing. As a first step, we test if proxies for limits-to-arbitrage and informational frictions are possible determinants of $\mathbb{E}$[MP]. Through panel regressions we find that higher idiosyncratic volatility, less shorting supply, and greater informational frictions contribute to higher

absolute mispricing. Furthermore, we find that additional shorting supply corresponds to a lower propensity of an overvaluation, while higher informational frictions impose the opposite. In line with the arguments developed by Baker and Wurgler (2006), we find that the realized mispricing spread between under- and overvalued stocks is amplified after periods of high investor sentiment. The amplification is driven almost exclusively by a more severe overvaluation (Stambaugh, Yu, and Yuan, 2012). We also find that earnings announcements play a pivotal role in correcting the mispricing of individual stocks (Engelberg et al., 2018) but that their impact is significantly smaller than previous studies suggest. Instead, our findings advocate that other mispricing proxies also pick up on compensation for common sources of variation.

Furthermore, we investigate the persistence of mispricing. On average, we find that both an over- and an undervaluation are corrected at a similar pace. Once we condition on the prevalence of short-selling constraints, however, we find that the overvaluation of stocks with limited short-selling supply persists substantially longer, and that much of their mispricing persists even after a full year. Smaller and illiquid firms, as well as firms that attract more retail interest also remain mispriced for a longer period of time. Higher idiosyncratic volatility, which is often used as a proxy for the costs of engaging in arbitrage activity, does not necessarily lead to a more persistent mispricing but the initial mispricing is significantly elevated. Finally, we revisit the recent evidence provided by van Binsbergen et al. (2023) and investigate which characteristic-sorted anomalies further exacerbate and which ones resolve stock-level mispricing, for which we use our mispricing proxy $\mathbb{E}[\text{MP}]$. Consistent with van Binsbergen et al. (2023), we find that some momentum-based and some profitability-related anomalies further exacerbate mispricing. Most size- and quality-based anomalies in turn align market prices with fair values. However, we find that this evidence is highly dependent on the information used to formulate the anomaly strategy and that there is substantial heterogeneity among the 13

characteristic groups identified by Jensen, Kelly, and Pedersen (2023).

Our third contribution is a detailed comparison of $\mathbb{E}[MP]$ with established proxies for mispricing. We consider the mispricing score by Stambaugh et al. (2015) (SYY) and the anomaly score "Net" by Engelberg et al. (2018) as competing measures. Both approaches rely on the intuition that some firm-level characteristics inform us about whether a stock is mispriced relative to the current cross-section. $\mathbb{E}[MP]$ is grounded in economic theory through the stochastic discount factor and allows for a directed assessment of mispricing and its quantification. In factor spanning regressions, we show that SYY and Net are unable to discern over- from undervalued stocks and that the resulting spread in excess returns can be explained by various factor models. Furthermore, the spread in realized mispricing, while significant for some of the factor models considered, has the wrong sign for both SYY and Net. $\mathbb{E}[MP]$ shows the correct sign for both return components and its return patterns are unexplained by the candidate factor models. Firm-level cross-sectional regressions confirm this evidence: $\mathbb{E}[MP]$ stands out as the sole mispricing measure with a consistent relationship with both excess returns and the mispricing return component, even after we control for various established firm characteristics (predictors) as well as SYY and Net. Independent double sorts between $\mathbb{E}[MP]$ and SYY or Net shed light on the underlying reason for this. Within each SYY or Net quintile, the mispricing spread uncovered by $\mathbb{E}[MP]$ is economically large and highly significant. In contrast, both SYY and Net cannot distinguish over- from undervalued stocks in the double sorts.

Finally, we provide a brief discussion of the aggregate implications of mispricing. We compute the share of the U.S. market capitalization that we currently expect to be overvalued and find that it fluctuates with the business cycle. It is higher in economic expansions and contracts during recessions. We also find that the two most recent decades show a smaller share of overvalued capital than the 90s. We regress the overvaluation share on various macroeconomic indicators and find that a one standard deviation increase

in the intermediary capital ratio (He, Kelly, and Manela, 2017) is accompanied by a 0.22 standard deviations increase in the overvaluation share, consistent with a greater willingness by intermediaries to lend capital. Higher market-wide net share issuance also amplifies the overvaluation share. This is in line with managers strategically exploiting elevated share prices (McLean, Pontiff, and Reilly, 2020). A higher Treasury bill rate, as a benchmark for the costs of short-selling, is also related to a higher overvaluation share, whereas greater uncertainty, which we measure with the CBOE's volatility index (VIX), translates to a lower overvaluation share. Finally, we show that the Chicago Fed National Activity Index, as a leading macroeconomic indicator, is positively related to the overvaluation share.

The remainder of the paper is organized as follows. Section 2 introduces a novel measure of firm-level expected mispricing. Section 3 first describes the data and key variables, and then presents the main empirical results. Section 4 investigates the economic underpinnings of the stock return predictability due to expected mispricing. Section 5 compares the empirical performance of our newly proposed measure of mispricing with existing measures. Section 6 provides economy-wide implications of the aggregate measure of mispricing. Section 7 concludes the paper.

2. Expected Mispricing

In this section, we introduce a novel measure of firm-level mispricing, which is grounded in economic theory. For this, we need two ingredients: first, a suitable factor model to obtain risk-adjusted returns of individual stocks. The return component unexplained by exposure to common factors constitutes a deviation from the stock's fundamental value. As the second ingredient, we need a method that is capable of predicting this mispricing. We leverage recent advances in machine learning to come up with our new measure of

expected mispricing, $\mathbb{E}[\text{MP}]$.

We start from a general definition of mispricing. For this, let m_{t+1} denote the stochastic discount factor. Under the assumption of "no arbitrage", the fair price of stock i with next period's payoff $X_{i,t+1}$ is given by the Euler equation:

$$P^*_{i,t} = \mathbb{E}_t \left[m_{t+1} X_{i,t+1} \right]. \tag{1}$$

A stock is mispriced whenever the currently observed market price, $P_{i,t}$, deviates from the true price, $P^*_{i,t}$. If $P_{i,t} > P^*_{i,t}$, stock i is overvalued, and if $P_{i,t} < P^*_{i,t}$ it is undervalued. This definition requires the specification of a suitable stochastic discount factor m_{t+1}. For this, we first express Eq. (1) in terms of excess returns:

$$\mathbb{E}_t \left[m_{t+1} r_{i,t+1} \right] = 0, \tag{2}$$

which can be rearranged to

$$\mathbb{E}_t \left[r_{i,t+1} \right] = \underbrace{\frac{Cov_t(m_{t+1}, r_{i,t+1})}{Var_t(m_{t+1})}}_{\beta_{i,t}} \underbrace{\left(- \frac{Var_t(m_{t+1})}{\mathbb{E}_t[m_{t+1}]} \right)}_{\lambda_t}. \tag{3}$$

When m_{t+1} is linear in factors F_{t+1}, we obtain the familiar factor representation of excess returns:

$$r_{i,t+1} = \beta_{i,t} F_{t+1} + \varepsilon_{i,t+1}, \tag{4}$$

with $\mathbb{E}_t[\varepsilon_{i,t+1} F_{t+1}] = 0$, and $\mathbb{E}_t[F_{t+1}] = \lambda_t$. To obtain a measure of stock-level mispricing, we need to specify a stochastic discount factor m_{t+1}, which is equivalent to specifying a suitable factor model.

We of course do not know the "true" factor model. In this paper, we therefore resort to *instrumented principal component analysis* (IPCA) by Kelly et al. (2019) as a statis-

tically optimized factor representation derived from the data. IPCA has been shown to significantly improve upon existing observable factor models by leveraging the information embedded in realized stock returns, without resorting to ex-ante knowledge about which factors may perform in a given period. In IPCA, firm characteristics $z_{i,t}$ are used to instrument for heterogeneity in stock i's sensitivities $\beta_{i,t}$ to K latent factors F_{t+1}:

$$r_{i,t+1} = \underbrace{\beta_{i,t} F_{t+1}}_{\text{Common Component}} + \underbrace{\varepsilon_{i,t+1}}_{\text{Mispricing=MP}} \quad \text{with} \quad \beta_{i,t} = z'_{i,t}\Gamma_\beta, \tag{5}$$

Following Kelly et al. (2019) who show that upwards of five factors are sufficient to drive out the explanatory power of a potential alpha ($\alpha_{i,t}$), we use the "restricted" IPCA specification that prevents characteristics to drive alphas. We thus allow for the common factors to pick up on as much common variation as possible and focus on mispricing as return variation above and beyond of what can be explained through common factor exposure. Note that this does not rule out the existence of mispricing per se. Kelly et al. (2019) argue that the restricted version "allows for some mispricing, as long as mispricings are truly idiosyncratic". Accordingly, a stock is mispriced whenever its return is poorly explained by the factor structure in Eq. (5). The residual return component $\varepsilon_{i,t+1}$ informs us about this. A negative MP shows that stock i's realized excess return is too small relative to the fair (expected) rate of return implied by the factor model. Assuming that mispricing is resolved over time, this is in line with an overvaluation. Conversely, a positive MP represents an undervaluation of stock i. In Section 4.4 we validate the assumption that on average mispricing disappears over time. We acknowledge the fact that in turn a negative (positive) MP could mean that an overvaluation (undervaluation) has build up and follow van Binsbergen et al. (2023) to classify cross-sectional asset pricing anomalies into those that exacerbate mispricing and those that resolve it in Section 4.5.

To showcase that IPCA outperforms other benchmark models in explaining the cross-

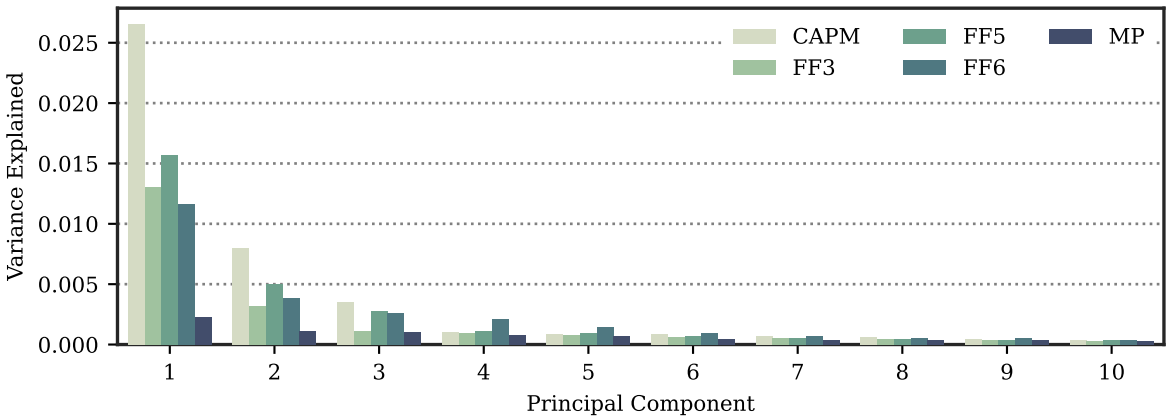


Fig. 2. Principal Components after Factor Adjustments.

The figure shows the variance explained by the first ten principal components across 25 portfolios sorted independently by size and book-to-market. Portfolios are formed by value-weighting the residual return component relative to five different candidate factor models. Next to the mispricing relative to IPCA (MP), we consider the CAPM, the Fama and French (1993) three-factor (FF3), the Fama and French (2015) five-factor (FF5) and the five-factor model augmented with momentum (Carhart, 1997, FF6). The figure displays the out-of-sample results for the period from January 1993 to December 2022.

section of stock returns, we form 25 value-weighted portfolios of stocks sorted independently on firms' size and book-to-market ratios. We regress each portfolio's returns on several candidate factor models obtained from Kenneth French's website, i.e., the CAPM, the Fama and French (1993) three-factor (FF3) model, the Fama and French (2015) five-factor (FF5) model, and the Fama and French (2018) six-factor (FF6) model that augments the FF5 model with the momentum factor of Carhart (1997), and also include the IPCA specification in Eq. (5).[4] We then analyze the remaining factor structure in the residuals obtained after controlling for common variation through each of the five factor models. To do so, we extract principal components from these residuals and plot the amount of variation explained by the first ten principal components in Figure 2. Given that the 25 portfolios are constructed from the same firm-level characteristics as the factors in the Fama and French (1993) model, we would naturally expect the model to explain these returns well.

[4] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

In contrast, we find that IPCA outperforms the other factor models by an incredibly wide margin. For example, the first principal component still explains a variance of more than 2.5% of monthly stock returns for the CAPM. Whereas the multi-factor models perform considerably better, only IPCA manages to pick up on most of the common linear variation in the cross-section of stock returns. Consequently, we favor IPCA to account for common sources of variation.

With a suitable measure of mispricing at hand, we now turn to the question of how to predict this mispricing ex-ante. We use the same set of firm-level characteristics as in IPCA to express *expected mispricing* as

$$\mathbb{E}_t[\mathrm{MP}_{i,t+1}] = g(z_{i,t}), \tag{6}$$

where $g(\cdot)$ is a flexible function mapping today's characteristics to expected mispricing. Gu et al. (2020) show that this setup works well to predict future stock excess returns. Because we are stripping excess returns of their common linear variation with IPCA, we need to use a prediction method capable of incorporating non-linearities and interactions between input variables. We approximate $g(\cdot)$ using well-established machine learning estimators: (*i*) a feed-forward neural network with three hidden layers and the pyramidal layering scheme described in Gu et al. (2020); (*ii*) a gradient-boosted regression tree (Friedman, 2001, GBT); and (*iii*) a random forest (Breiman, 2001, RF). Tree-based prediction methods have shown incredible success on structured datasets, whereas Gu et al. (2020) show that neural networks outperform competing models for stock return prediction. To strike a balance between the benefits and drawbacks of the three methods, we form an equal-weighted ensemble, which has been shown to boost the predictive performance significantly (Rapach et al., 2010; Bali, Beckmeyer, Moerke, and Weigert,

2023):

$$\mathbb{E}_t[\text{MP}_{i,t+1}] = \frac{1}{3} \cdot \sum_{j=1}^{3} g_j(z_{i,t}). \tag{7}$$

Furthermore, we optimize the hyperparameters within each model type and use the average prediction of the eight best-performing models to construct model j's representation of $g_j(\cdot)$, which further robustifies the ensemble's performance.

3. Firm-level Expected Mispricing

Data and Model Estimation. We use the extensive dataset of monthly firm characteristics and stock returns studied by Jensen et al. (2023). We obtain the data from CRSP and Compustat for all stocks trading on NYSE, NASDAQ and AMEX. We focus on the 153 characteristics studied by Jensen et al. (2023) and start our sample in January 1973. Before that, many characteristics have limited availability. To deal with outliers and noise in the raw characteristics we rank-standardize the features between $[-0.5, 0.5]$ (Kelly et al., 2019).

The estimation of *expected mispricing* consists of two steps. First, we estimate mispricing MP relative to the IPCA benchmark model as in Eq. (5). To mimic the behavior of an investor planning to identify mispriced investment opportunities in the market, we carefully set up an estimation procedure that avoids the inclusion of forward-looking information. Specifically, for each month t, we use characteristics and stock returns up until month t to estimate Γ_β. We then hold Γ_β fixed over the next month to obtain factor realizations F_{t+1} out-of-sample. This way, the weights of each stock to construct the latent factors are known at the time of the investment decision at the end of month t. We use the extracted factor structure and stock-specific βs to strip out the common return component, which gives us the realized mispricing in the next month, $\text{MP}_{i,t+1}$.

The second step is to estimate the expected mispricing measure using firm-level characteristics (see Eq. 7). We select the first 70% of months up until month t to train the model's parameters and validate the model's hyperparameters in the remaining 30% of months. We then apply the trained model to obtain our expected mispricing measure $\mathbb{E}_t[\text{MP}_{i,t+1}]$. As re-estimating machine learning models is more costly than IPCA, we refit each of the machine learning estimators at the end of December of each year and hold the model fixed for the entire next year. For IPCA, we re-estimate the model parameters every month. Our first out-of-sample period starts in January 1993. We use the preceding 20 years of data – January 1973 to December 1992 – to estimate and validate the models. We then expand the training set by the months in 1993 and re-estimate the models for predictions in 1994. We repeat this process until all years are exhausted, with December of 2022 marking the end of our 30-year long out-of-sample period.[5]

Model Performance. We assess the model's performance in the 30-year long out-of-sample period $\mathcal{T}_3$ using the standard R^2 statistic:

$$R^2_{\text{OOS}} = 1 - \frac{\sum_{(i,t)\in\mathcal{T}_3}(\text{MP}_{i,t+1} - \mathbb{E}_t[\text{MP}_{i,t+1}])^2}{\sum_{(i,t)\in\mathcal{T}_3}\text{MP}^2_{i,t+1}}. \tag{8}$$

While R^2_{OOS} gives the *total* amount of variation that is explained by the model's predictions above a naive benchmark of zero, Han, He, Rapach, and Zhou (2023) propose a simple modification to focus on the amount of *cross-sectional* variation which can be explained:

$$R^2_{\text{OOS, XSEC}} = 1 - \frac{\sum_{(i,t)\in\mathcal{T}_3}[(\text{MP}_{i,t+1} - \overline{\text{MP}}_{t+1}) - (\mathbb{E}_t[\text{MP}_{i,t+1}] - \overline{\mathbb{E}_t[\text{MP}]}_{t+1})]^2}{\sum_{(i,t)\in\mathcal{T}_3}(\text{MP}_{i,t+1} - \overline{\text{MP}}_{t+1})^2}, \tag{9}$$

[5]We provide the code to replicate the empirical findings of this paper, as well as the prepared *expected mispricing* per firm between 1993 and 2022 on GitHub.

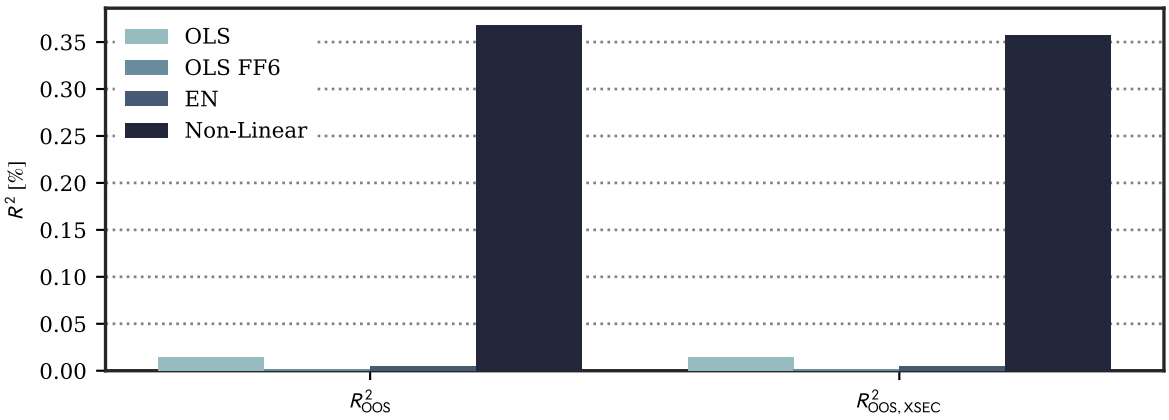


Fig. 3. Model Performance for Mispricing Prediction.

The figure shows the R^2 statistic of several models aiming to predict a stock's future mispricing. We show both the out-of-sample R^2 (Eq. (8)) and the cross-sectional out-of-sample $R2$ (Eq. (9)). We compare an ordinary least squares using all 153 firm-level characteristics (OLS), an ordinary least squares using just the Fama and French (2015) characteristics augmented with momentum (OLS FF6), an elastic net (EN), and our non-linear ensemble model (Non-Linear). R^2s are multiplied by 100. The figure displays the out-of-sample results for the period from January 1993 to December 2022.

where $\overline{\text{MP}}_{t+1}$ and $\overline{\mathbb{E}_t[\text{MP}]}_{t+1}$ denote the time-t cross-sectional average of realized (MP) and expected mispricing ($\mathbb{E}$[MP]), respectively. Results are shown in Figure 3.

As a sanity check, we also assess the predictive power for different linear models. This includes a simple OLS regression (OLS) on the entire set of 153 firm characteristics, an OLS regression using the characteristics used in the Fama and French (2015) factor model augmented by momentum (OLS FF6), and an elastic net model (Zou and Hastie, 2005, EN). If the latent factor structure in Eq. (5) is expressive enough, an attempt to estimate a linear function of stock characteristics to predict MP should be futile. This is indeed what we find: all linear models fail to predict future mispricing with R^2s close to zero. In contrast, our ensemble model denoted by "Non-Linear" has an out-of-sample R^2_{OOS} of 0.37%, which is roughly comparable to the 0.40% that Gu et al. (2020) find for the prediction of total excess stock returns. Our ensemble machine learning model does equally well in describing cross-sectional return spreads with a cross-sectional out-of-sample $R^2_{\text{OOS, XSEC}}$ of 0.36%. Overall, we document that our model can predict future

Table 1: Expected Mispricing-based Trading Strategy

The table shows performance statistics for decile portfolios sorted by our expected mispricing measure $\mathbb{E}[\text{MP}]$. We show the average expected mispricing $\mathbb{E}[\text{MP}]$, realized mispricing (MP), mispricing variation ($\sigma(\text{MP})$), the information ratio ($\text{MP}/\sigma(\text{MP})$), realized excess return (r) and the Sharpe ratio. We calculate HAC-robust standard errors using 12 lags and follow Lo (2002) in calculating the statistics of both the Sharpe and the information ratio. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	$\mathbb{E}[\text{MP}]$ [%]	MP [%]	$\sigma(\text{MP})$ [%]	IR	r [%]	SR
Overvalued	−1.23	−1.12	3.71	−1.04	−0.86	−0.28
2	−0.57	−0.40	2.50	−0.56	0.25	0.12
3	−0.36	−0.38	2.31	−0.57	0.43	0.31
4	−0.23	−0.08	2.20	−0.13	0.66	0.49
5	−0.12	0.02	2.17	0.04	0.70	0.56
6	−0.03	0.03	2.24	0.05	0.68	0.53
7	0.07	0.16	2.13	0.26	0.59	0.41
8	0.18	0.30	2.32	0.45	0.64	0.43
9	0.33	0.56	2.19	0.88	0.71	0.44
Undervalued	0.63	1.01	2.63	1.33	1.06	0.55
HL	1.86***	2.13***	4.49	1.64***	1.92***	0.89***

stock-level mispricing and that the incorporation of non-linearities are vital to its success.

Consider once more our hypothetical investor seeking to identify mispriced investment opportunities in the market. She may follow a simple strategy of sorting stocks into decile portfolios by $\mathbb{E}[\text{MP}]$ and documenting the realized return and mispricing spreads between the most under- and the most overvalued stocks. To limit the influence of microcraps, we weight stocks by their market capitalization. The results to this exercise are shown in Table 1.

Profitability. Our ensemble of machine learning models predicts an average MP spread of 1.86% per month, which is close to the realized high-minus-low MP spread of monthly 2.13%. This is remarkable as MP is the return component, which is unexplained by six latent factors. The average realized MP over the next month of the most overvalued stocks amounts to −1.12%, for the most undervalued stocks it amounts to 1.01.%. We find that the time-series variation of realized MP ($\sigma(MP)$) is comparable for most port-

folios but slightly elevated for the most overvalued stocks. This also shows up in the information ratio (IR), defined as MP/σ(MP). While the absolute realized mispricing is largest among the most overvalued stocks, we find that the resulting IR is less extreme at -1.04, compared to an IR of 1.33 for the most undervalued stocks. The IR of the high-minus-low spread portfolio amounts to a whopping 1.64, highlighting that our setup identifies significant investment opportunities which cannot be explained by stocks' exposure to common factors. The one-month-ahead average return spread and realized Sharpe ratio are also economically large at 1.92% per month and an annualized 0.89, respectively, and show that a significant portion of the expected mispricing component shows up in realized returns.

It is natural to assume that mispricing is more prevalent among smaller firms, which receive less media attention and capital from institutional investors. We therefore study the influence of firm size on the profitability of the high-minus-low spread portfolio sorted by $\mathbb{E}$[MP] in Appendix A. We continue to value-weight the stocks in each portfolio but now cap the weights at the xth market capitalization percentile in month t, $\min[MC_{i,t}, Q_t^x(MC)]$. This way, we limit the impact of megacaps, which potentially skew portfolio-level results (Jensen et al., 2023).

Consistent with the notion that mispricing affects smaller stocks more, we find the largest spreads in realized returns and mispricing when we weight stocks equally (capping at the zero'th market capitalization percentile). The spread in average realized mispricing increases from 2.13% per month in the case of value-weighting to 3.06% for equal weights; the spread in average realized returns increases even further to 3.78% per month. The information (IR) and Sharpe ratio (SR) also increase monotonically as we limit the impact of megacaps and top out at 3.4 and 2.9, respectively, when the stocks receive an equal weight in the over- and undervalued portfolios.

Table 2: Descriptive Statistics for Expected Mispricing-based Portfolios

The table shows descriptive statistics for decile portfolios sorted by our expected mispricing measure $\mathbb{E}[MP]$. We show average values for each portfolio for book-to-market (B2M), market capitalization (Size), idiosyncratic volatility (Ang et al., 2006, IVOL), momentum (Mom), illiquidity (Illiq), profitability (Prof), max return (MAX), price (PRC), investment (Inv) and analyst earnings forecast dispersion (DISP) following Diether et al. (2002). Note that all characteristics have been cross-sectionally rank-standardized to the interval $[-0.5, 0.5]$. We calculate HAC robust standard errors using 12 lags. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	B2M	Size	IVOL	Mom	Illiq	Prof	MAX	PRC	Inv	DISP
Overvalued	0.01	−0.23	0.28	−0.17	0.18	−0.23	0.26	−0.26	−0.06	0.17
2	0.04	−0.14	0.05	−0.03	0.13	−0.12	0.02	−0.12	−0.06	0.09
3	0.04	−0.05	−0.04	0.01	0.06	−0.04	−0.06	−0.03	−0.05	0.02
4	0.02	0.02	−0.09	0.03	−0.01	0.02	−0.10	0.03	−0.03	−0.02
5	0.01	0.06	−0.10	0.03	−0.05	0.05	−0.10	0.06	−0.03	−0.04
6	0.00	0.08	−0.10	0.03	−0.07	0.07	−0.09	0.08	−0.02	−0.05
7	−0.00	0.09	−0.09	0.02	−0.08	0.08	−0.07	0.08	−0.02	−0.04
8	−0.01	0.10	−0.06	0.02	−0.08	0.08	−0.03	0.09	−0.02	−0.03
9	−0.05	0.10	−0.01	0.03	−0.09	0.08	0.03	0.09	−0.01	−0.02
Undervalued	−0.06	−0.02	0.15	0.01	0.01	−0.02	0.16	−0.02	−0.03	0.03
HL	-0.08^{***}	0.21^{***}	-0.13^{***}	0.17^{***}	-0.17^{***}	0.21^{***}	-0.11^{***}	0.24^{***}	0.03^{***}	-0.14^{***}

Which Stocks are Mispriced? To determine which stocks are characterized as over- or undervalued, Table 2 provides information on a selected set of firm characteristics for the decile portfolios of stocks sorted by $\mathbb{E}[MP]$ and the resulting high-minus-low spread portfolio. We include each firms' book-to-market (B2M), market capitalization (Size), idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang, 2006, IVOL), momentum (Mom), illiquidity (Illiq), profitability (Prof), max daily return over the past month (MAX), price (PRC), investment (Inv), and the analyst earnings forecast dispersion (DISP) following Diether et al. (2002). Importantly, we cross-sectionally rank-standardize characteristics to the interval $[-0.5, 0.5]$ to make the values comparable over time and with a changing cross-sectional breadth. This also eases the interpretation. For example, the size spread of 0.21 between under- and overvalued stocks suggests that undervalued stocks are, on average, 21 cross-sectional percentiles larger than overvalued stocks.

The most glaring differences in the characteristics of the most under- and the most overvalued stocks are in the stocks' market capitalization (spread of 0.21), the stocks'

profitability (0.21) and price (0.24). This suggests that overvalued stocks tend to be significantly smaller, less profitable, and have smaller nominal share prices. Profitability (Prof) and investment (inv) are also part of the selection variables of the Stambaugh et al. (2015) mispricing measure. Our measure of *expected mispricing*, which is grounded in economic theory, finds a large spread in profitability but an economically negligible spread in the investment behavior of under- vs. overvalued firms. The spread in the book-to-market ratio is also small in comparison. That overvalued stocks have smaller nominal prices and larger MAX returns aligns well with the literature. Birru and Wang (2016) show that investors overestimate the skewness of low-priced stocks and Bali et al. (2011) find that investors' appetite for positive skewness is well approximated with the maximum daily return in the past month. The resulting demand by bias-afflicted investors leads to an overvaluation. We furthermore find that overvalued stocks tend to be less liquid (Amihud, 2002) and in the lower half of the momentum distribution. In addition, we find that the idiosyncratic volatility (IVOL), estimated following Ang et al. (2006), is high for both under- and overvalued stocks, consistent with the finding of Stambaugh et al. (2015) that IVOL is a proxy for arbitrage costs. IVOL is most elevated for overvalued stocks with an average rank of 0.28 in a range of -0.5 and 0.5. Finally, we find that overvalued stocks tend to have a higher dispersion in analyst earnings forecasts when compared to undervalued stocks. The dispersion measure (DISP), which follows Diether et al. (2002), averages 0.17 for overvalued stocks, which is 0.14 percentiles higher than the value for undervalued stocks. This finding is in line with the argument of Miller (1977) that stocks with higher investor disagreement are more likely to be overpriced, and should hence generate lower future returns. In Miller's framework, future underperformance of overvalued stocks with high investor disagreement is justified by short-selling constraints that sideline more pessimistic investors, such that prices are predominantly determined by optimistic investors.

4. The Economics of Mispricing

In this section, we investigate the economic underpinnings of the stock return predictability due to expected mispricing. We relate our measure of *expected mispricing* to various proxies for limits to arbitrage, study the impact of investor sentiment on the predictability of mispriced investment opportunities, as well as the differential resolution of mispricing around earnings announcements, examine the persistence of mispricing in the cross-section of individual stocks, and investigate which cross-sectional return anomalies resolve and which ones exacerbate the stock-level mispricing.

4.1. Limits-to-Arbitrage

Determinants of Mispricing. We first regress $\mathbb{E}[\text{MP}]$ on several proxies for impediments to arbitrage activity. We include a stock's idiosyncratic volatility as a measure of arbitrage costs (Stambaugh et al., 2012). In line with Table 2, we expect stocks with higher idiosyncratic volatility to be more likely to be mispriced. A similar argument applies to institutional ownership, which Nagel (2005) proposes as a proxy for short-selling supply. Less short-selling supply limits the activity of arbitrageurs which in turn translates to a higher prevalence and more severe mispricing, especially in the form of overpricing. As another proxy for the activity of short-sellers, we include a stock's current short interest, as a fraction of its market capitalization. Finally, we follow Atilgan, Bali, Demirtas, and Gunaydin (2020) and construct an informational frictions index (IF) to capture the various facets of limits-to-arbitrage. We combine the informational content of multiple limits-to-arbitrage proxies into a composite score: we include a stock's idiosyncratic volatility, institutional ownership share, its market capitalization, firm age, Amihud (2002)'s illiquidity measure, and the number of analysts covering a firm (Zhang, 2006). To construct the arbitrage cost index, we sort stocks in increasing order based on

Table 3: Determinants of Expected Mispricing

The table shows the output of a panel regression of both the level $\mathbb{E}[MP]$ and the absolute level $|\mathbb{E}[MP]|$ of expected mispricing on several proxies. These include each stock's aggregate short interest (Short Sell), institutional ownership (IO) as a proxy for short-selling supply (Nagel, 2005), idiosyncratic volatility (IVOL) as a proxy for arbitrage risk, and an informational frictions index (IF) which follows Atilgan et al. (2020). Expected mispricing is given in percentage values. The remaining variables are normalized. We include day and entity fixed effects and cluster standard errors by month and entity. The table displays the out-of-sample results for IVOL for the period from January 1993 to December 2022. Data availability for the other variables restricts the sample to end in March 2021, and for IF to start in January 2000.

	Expected Mispricing				Absolute Expected Mispricing			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IVOL	−0.129				0.293			
	[−18.35]				[68.92]			
IO		0.046				−0.098		
		[7.73]				[−21.64]		
Short Sell			0.056				−0.038	
			[11.69]				[−10.66]	
IF				−0.149				0.261
				[−18.34]				[43.85]
FixedEffects	T+E	T+E	T+E	T+E	T+E	T+E	T+E	T+E
R^2	2.10	0.09	0.50	1.02	17.59	0.71	0.39	5.27

their idiosyncratic volatility and Amihud (2002) illiquidity. Similarly, we sort stocks in decreasing order based on their level of institutional ownership, analyst coverage, size, and age since lower values of these variables indicate higher arbitrage costs. Each stock is given the corresponding score of its quintile rank for each variable. Finally, the index on the stock level is the sum of the six scores, ranging from 6 to 30. A higher value implies tighter limits-to-arbitrage and a larger degree of mispricing. We also include month fixed effects and entity fixed effects to account for market-level effects and unobserved firm characteristics.

The results of regressing $\mathbb{E}[MP]$ on the proxies for limits-to-arbitrage are provided in Table 3. Consistent with the characteristic-based evidence presented in Table 2, we find that a higher IVOL is related to a larger expected overvaluation of these stocks. When performing the same regression using $|\mathbb{E}[MP]|$ as the dependent variable, we find that a one standard deviation increase in IVOL corresponds to an increase in absolute

expected mispricing of 0.293% over the next month. This is consistent with the notion that idiosyncratic volatility is a rough proxy for the costs of arbitrage activity (Stambaugh et al., 2015). Additional shorting supply lowers the chance that the target stock is overvalued. A one standard deviation decrease in institutional ownership (or higher retail ownership) corresponds to 4.6 bps larger expected overpricing. A one standard deviation increase in the short-selling activity elevates $\mathbb{E}[\text{MP}]$ by 5.6 bps. Looking at the absolute expected mispricing, we find that additional shorting supply and activity decrease the overall magnitude of both an under- and an overvaluation. Finally, we find that a higher level of informational frictions (IF) corresponds to a greater likelihood of an overvaluation and in general a higher propensity of the target stock to be mispriced ($|\mathbb{E}[\text{MP}]|$). All coefficients are highly statistically significant using standard errors clustered by month and entity. While these effects are large, they also show that much of the variation in $\mathbb{E}[\text{MP}]$ is unexplained by the considered proxies. $\mathbb{E}[\text{MP}]$ is an intricate mixture of proxies for limits-to-arbitrage and other firm characteristics, with potentially complex non-linearities and interactions between them.

Realized Mispricing. *Expected mispricing* is related to various proxies for limits to arbitrage (Table 3) and there are substantial characteristic-level differences between the most over- and the most undervalued firms (see Table 2). We now perform independent quintile sorts on $\mathbb{E}[\text{MP}]$ and idiosyncratic volatility, institutional ownership share, and the composite informational frictions index, to assess if systematic differences in *realized* mispricing exist for groups of stocks that we expect to differ in how easily arbitrage activity is conducted and information impounded into market prices. We record the value-weighted realized mispricing for each portfolio over the next month.

Table 3 reports that IVOL is an important determinant of the absolute level of $\mathbb{E}[\text{MP}]$ and we now show in Table 4 that it is also an important conditioning variable for real-

Table 4: Double Sort on Expected Mispricing and Idiosyncratic Volatility.

The table reports value-weighted averages of *realized mispricing* for 5x5 portfolios independently sorted on our measure of expected mispricing $\mathbb{E}[MP]$ and the stock's idiosyncratic volatility (IVOL). That is, in each month t we calculate our measure of expected mispricing and the stock's IVOL, form portfolios based on these measures and report the realized mispricing for the next month $t+1$. We calculate HAC-robust standard errors using 12 lags. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	Expected Mispricing					
IVOL	Overvalued	2	3	4	Undervalued	HL
LOW	-0.28^{**}	-0.07	0.18	0.38^{**}	0.52^{***}	0.80^{***}
2	-0.66^{***}	-0.42^{**}	-0.22	0.14	0.66^{***}	1.32^{***}
3	-0.50^{**}	-0.51^{***}	-0.20	0.05	0.81^{***}	1.30^{***}
4	-1.04^{***}	-0.67^{***}	-0.40^{*}	0.02	0.98^{***}	2.02^{***}
HIGH	-1.10^{***}	-0.63^{**}	-0.07	0.65^{**}	1.43^{***}	2.53^{***}
HL	-0.82^{***}	-0.56	-0.24	0.28	0.92^{**}	1.73^{***}

ized mispricing over the next month. The high-minus-low spread in realized mispricing amounts to 0.80% for low-IVOL stocks and increases nearly monotonically to 2.53% per month for stocks in the highest IVOL quintile. The difference-in-difference mispricing spread stands at 1.73% and is highly significant. In line with the theory developed in Stambaugh et al. (2015), we find that higher IVOL increases the magnitude of both an under- *and* an overvaluation. While the most overvalued stocks with a low IVOL on average realize a mispricing of -0.28% in the next month, overvalued stocks in the highest IVOL bucket average -1.10% – a near fourfold increase. Among undervalued stocks, we find a similar pattern: low IVOL stocks that are undervalued have a realized mispricing return component of 0.52% over the next month, compared to 1.43% for undervalued stocks in the high IVOL bucket.

The results for the institutional ownership share in Table 5 also align well with earlier studies. Consistent with the idea that low institutional interest translates to less shorting supply (Nagel, 2005), we find higher levels of over- and undervaluation in the lowest IO bucket. The high-minus-low spread between under- and overvalued stocks largely held by retail investors (low IO) amounts to 2.26%. The spread down to the highest IO

Table 5: Double Sort on Expected Mispricing and Institutional Ownership.

The table reports value-weighted averages of *realized mispricing* for 5x5 portfolios independently sorted on our measure of expected mispricing $\mathbb{E}[MP]$ and the stock's share of institutional ownership (IO). That is, in each month t we calculate our measure of expected mispricing and the stocks IO, form portfolios based on these measures and report the realized mispricing for the next month $t+1$. We calculate HAC-robust standard errors using 12 lags. The table displays the out-of-sample results for the period from January 1993 to May 2021.

	Expected Mispricing					
IO	Overvalued	2	3	4	Undervalued	HL
LOW	-1.39^{***}	-0.53^{***}	-0.06	0.20	0.87^{***}	2.26^{***}
2	-0.58^{***}	-0.17	-0.12	0.50^{***}	1.04^{***}	1.62^{***}
3	-0.46^{*}	-0.17	0.18	0.28^{*}	0.74^{***}	1.20^{***}
4	-0.34	-0.16	0.07	0.27	0.72^{***}	1.06^{***}
HIGH	-0.24	-0.15	-0.04	0.11	0.69^{***}	0.93^{***}
HL	1.15^{***}	0.37^{**}	0.02	-0.10	-0.19	-1.33^{***}

bucket drops down monotonically to 0.93% but remains highly significant for all IO quintiles. Of course, less shorting supply should disproportionately impact overvalued firms. Correcting an overvaluation requires engaging in short-selling which becomes more costly or infeasible with less shorting supply. We find that most of the difference in the realized mispricing spreads between under- and overvalued stocks can be traced back to differences in the realized mispricing of overvalued stocks. The stocks with the lowest institutional ownership average a realized mispricing component of -1.39% compared to just -0.24% per month for the stocks in the highest IO bucket.

Finally, we repeat the analysis for the composite informational frictions (IF) index. How we construct the IF index is described earlier. Intuitively, higher informational frictions should translate into a higher magnitude of mispricing. We document results in line with this conjecture in Table 6. The realized mispricing spread between under- and overvalued stocks amounts to 0.88% among stocks with a low informational frictions index and rises to 1.83% per month for stocks in the highest IF bucket. The difference-in-difference portfolio shows that the difference of 0.95% is both economically and statistically significant. Absolute mispricing is thus generally stronger among stocks

Table 6: Double Sort on Expected Mispricing and Informational Frictions Index.

The table reports value-weighted averages of *realized mispricing* for 5x5 portfolios independently sorted on our measure of expected mispricing $\mathbb{E}$[MP] and the stock's informational frictions index (IF). That is, in each month t we calculate our measure of expected mispricing and the stock's IF, form portfolios based on these measures and report the realized mispricing for the next month $t+1$. We calculate HAC-robust standard errors using 12 lags. The table displays the out-of-sample results for the period from January 2000 to May 2021.

	Expected Mispricing					
IF Index	Overvalued	2	3	4	Undervalued	HL
LOW	-0.50^{**}	-0.15	-0.10	0.15	0.38^{***}	0.88^{***}
2	-0.53^{*}	-0.33^{**}	-0.12	0.03	0.71^{***}	1.24^{***}
3	-0.60^{**}	-0.34^{***}	0.07	0.14	0.49^{***}	1.09^{***}
4	-0.64^{***}	-0.44^{***}	-0.15	0.22^{*}	0.61^{***}	1.25^{***}
HIGH	-1.05^{***}	-0.26^{*}	0.55^{**}	-0.01	0.78^{***}	1.83^{***}
HL	-0.55^{*}	-0.11	0.64^{**}	-0.17	0.40	0.95^{***}

with higher informational frictions.

4.2. Sentiment

There is an intricate connection between the degree and prevalence of mispricing in the cross-section of stock returns and investor sentiment. Baker and Wurgler (2006) develop a theory in which sentiment predominantly impacts the stocks with higher limits-to-arbitrage and higher informational frictions, for which they find empirical support in the data. Stambaugh et al. (2012) extend this idea and argue that short-selling constraints are the root cause for why mispricing is not corrected immediately. Accordingly, an overvaluation should be most present after periods of high investor sentiment. The authors show that the returns of various anomaly portfolios concentrate after these periods. We now study the influence of investor sentiment on our measure of *expected mispricing*. From the theories outlined above, we can make two predictions: (*i*) the spread in realized mispricing between under- and overvalued stocks should be more pronounced after periods of high investor sentiment; and (*ii*) that the differences in this spread should be

Table 7: Impact of Sentiment

The table reports the average realized mispricing for decile portfolios sorted on our measure of expected mispricing $\mathbb{E}[MP]$. That is, in each month t we calculate our measure of expected mispricing, form value-weighted decile portfolios and report the realized mispricing for the next month $t+1$. We show time-series averages conditional on sentiment (SENT) (Baker and Wurgler, 2006) being above (HIGH) or below (LOW) its median value. We calculate HAC-robust standard errors using 12 lags. The table displays the out-of-sample results for the period from January 1993 to June 2022.

	Expected Mispricing										
SENT	Overvalued	2	3	4	5	6	7	8	9	Undervalued	HL
LOW	−0.63	−0.32	−0.25	0.08	0.18	0.13	0.39	0.48	0.54	1.00	1.63
	[−2.61]	[−1.96]	[−1.53]	[0.55]	[1.16]	[0.84]	[2.40]	[2.59]	[3.68]	[3.59]	[3.81]
HIGH	−1.57	−0.48	−0.47	−0.26	−0.07	−0.00	−0.03	0.18	0.62	1.05	2.62
	[−4.35]	[−2.17]	[−1.92]	[−1.22]	[−0.32]	[−0.02]	[−0.11]	[0.74]	[2.91]	[4.11]	[6.91]
HL	−0.94	−0.17	−0.22	−0.34	−0.25	−0.14	−0.42	−0.30	0.08	0.05	0.99
	[−2.39]	[−0.62]	[−0.91]	[−1.47]	[−1.09]	[−0.57]	[−1.86]	[−1.20]	[0.34]	[0.17]	[2.07]

driven by overvalued stocks, which are most impacted by the high sentiment.

We partition our sample into periods of high and low investor sentiment using the in-sample median of the Baker and Wurgler (2006) sentiment index. Within each subsample, we sort stocks into deciles by $\mathbb{E}[MP]$ and record the value-weighted realized mispricing of each portfolio in the following month. The results in Table 7 are in line with our predictions. First, we find that the realized mispricing spread is larger after periods of high investor sentiment (2.62% per month) compared to that after periods of low investor sentiment (1.63%). The difference-in-difference mispricing spread of 0.99% is economically large and statistically significant. Second, as we would expect, the difference in this spread is primarily driven by changes in the realized mispricing of overvalued stocks. While the realized mispricing averages at −0.63% after periods of low sentiment, it climbs to −1.57% after periods of high sentiment. This difference is large and significant for overvalued stocks; −0.94% per month (t-stat.=−2.39), whereas we find no difference between periods of low and high investor sentiment for undervalued stocks. All the same, $\mathbb{E}[MP]$ also uncovers a subset of stocks with an economically and statistically significant undervaluation.

Table 8: Anomaly Returns on Earnings Days

The table reports the result of a regression of daily returns on different measures of mispricing and their interaction with an earnings announcement dummy. We consider a symmetric three-day window around the official announcement day. For better comparability, all scores are cross-sectionally standardized to have unit variance and a mean zero. Daily returns are expressed in percentage points. We compare the results using our measure of expected mispricing $\mathbb{E}[MP]$ with the results for the anomaly score Net of Engelberg et al. (2018). The table replicates and extends Table III of Engelberg et al. (2018). We include day fixed effects and cluster standard errors by day and entity. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	Daily Excess Return			
	(1)	(2)	(3)	(4)
$\mathbb{E}[MP]$	0.094		0.094	0.093
	[27.75]		[27.96]	[27.53]
$\mathbb{E}[MP] \times EA$	0.062			0.035
	[4.89]			[2.81]
Net		0.020	0.007	0.007
		[4.98]	[1.73]	[1.75]
Net × EA		0.131	0.131	0.126
		[13.26]	[13.28]	[13.02]
EA	0.003	0.006	0.002	0.001
	[0.32]	[0.60]	[0.23]	[0.12]
FixedEffects	T	T	T	T
R^2	0.038	0.005	0.040	0.040

4.3. Earnings Announcements

The literature has also produced an interesting link between the prevalence of mispricing and the return patterns around earnings announcement days. Engelberg et al. (2018) argue that biased expectations are corrected around earnings announcements and show that anomaly returns are six times larger surrounding these announcements. For this, they rely on a simple anomaly score, which counts how often a stock is in the long leg of the arbitrage portfolio minus how often it is in the short leg of the arbitrage portfolio of various characteristic-based return anomalies. In Table 8, we investigate the resolution of mispricing around earnings announcements using our measure of *expected mispricing* and replicate the main result presented in Table III of Engelberg et al. (2018).

We regress daily returns on $\mathbb{E}[MP]$. For replication purposes, we recreate their anomaly

score "Net" using the 153 firm characteristics of Jensen et al. (2023) and regress daily returns also on Net. We additionally interact both measures of mispricing with an earnings announcement dummy, which captures the additional return response in a symmetric three-day window around the earnings announcement day. As in Engelberg et al. (2018), we add time fixed effects as controls. In column (2) of Table 8, we show that our extended Net score replicates the main result of Engelberg et al. (2018) almost one-to-one. Net's return impact is amplified more than six-fold on earnings announcement days, comparing the slope coefficients on Net (0.020) and Net $\times$ EA (0.131). In contrast to this, we find a much smaller additional impact of $\mathbb{E}[\text{MP}]$ on earnings announcements. While we do find that the impact of mispricing is amplified on earnings announcement days, in line with their theory of the correction of biased expectations, we find that the degree to which this happens is severely muted. Column (3) of Table 8 shows that on non-earnings announcement days, a one standard deviation increase in $\mathbb{E}[\text{MP}]$ translates to 9.4bps larger returns. As reported in column (1) of Table 8, on earnings announcement days this effect amplifies to 15.6bps (9.4 + 6.2) – a highly significant increase of 66%. Including both measures simultaneously leads to a similar conclusion. Interestingly, $\mathbb{E}[\text{MP}]$ subsumes the explanatory power of Net on all days, except for on stock-specific earnings announcement days. Overall, these results suggest that the anomaly score Net at least in part picks up on systematic compensation earned around earnings announcements but that earnings announcements continue to be an important event for the correction of the mispricing of individual stocks.

4.4. Persistence of Mispricing

We have so far examined the realization of mispricing relative to the IPCA model specified in Eq. (7) over the next month. We now investigate the persistence of mispricing. This

also allows us to test the assumption that any mispricing should resolve over time, which in turn validates our modeling approach. For this, we sort stocks into decile portfolios based on $\mathbb{E}[\text{MP}]$, keep the portfolio allocation fixed and study each portfolio's realized mispricing over the next $t + \Delta t$ months. The results for the 10% most under- and the 10% most overvalued stocks are shown in Figure 4. In line with the idea that mispricing resolves over time, we see a gradual decline in future realized mispricing in the subsequent months. The magnitude of an over- and an undervaluation is corrected at a similar pace. When looking at the statistical significance of future realized mispricing, however, we find that it takes an average of $\Delta t = 3$ months for an overpricing and a total of $\Delta t = 7$ months for an underpricing to be corrected.

This result may at first seem at odds with conventional wisdom in that an overvaluation should be purged from the market more slowly in the presence of costly short-selling (Stambaugh et al., 2015). To investigate the role of *binding* short-selling constraints, we investigate the persistence patterns of stocks that are more likely to be short-sale constrained: we follow Nagel (2005) and sort the stocks within each $\mathbb{E}[\text{MP}]$ decile port-

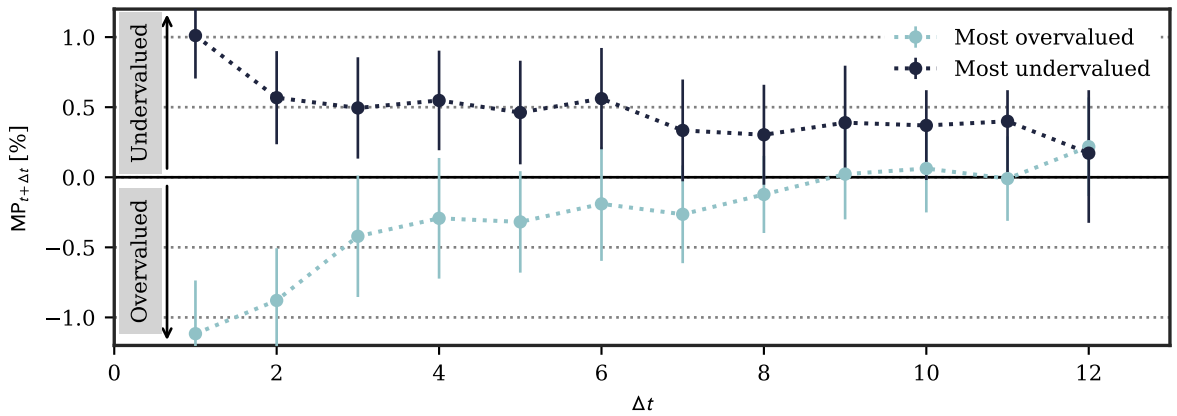


Fig. 4. Mispricing Persistence

The figure shows the decay in mispricing over time. In month $t = 0$ we sort stocks into deciles by our expected mispricing measure $\mathbb{E}[\text{MP}]$ and record the future realized mispricing in months $t + 1$ to $t + 12$ keeping the portfolio allocation fixed. We plot the results for the deciles of most undervalued and most overvalued stocks. Error bars denote the 5% confidence levels. The figure displays the out-of-sample results for the period from January 1993 to December 2022.

folio into terciles by their institutional ownership share. Since institutional investors commonly facilitate the short-selling of stocks, we expect more persistent overvaluation among stocks with high retail ownership. The results in Figure E.1 support this hypothesis; i.e., overvalued stocks with more binding short-sale constraints remain overvalued for a longer period of time. Specifically, we find that their overvaluation persists even after 12 months. In contrast, the overpricing of stocks with a heavy presence of institutional investors is corrected in as little as $\Delta t = 2$ months. We find only small differences in the persistence of mispricing among undervalued stocks with a high and a low institutional ownership share.

In Table 9, we perform a similar sorting exercise, showing the differential persistence in mispricing of the high-minus-low $\mathbb{E}$[MP] portfolios, where each $\mathbb{E}$[MP] portfolio is further subdivided into terciles by several proxies for limits-to-arbitrage and the interest of less sophisticated (retail) investors. As before, we include each stock's institutional ownership share (IO), its market capitalization (Size), nominal price (PRC), Amihud (2002) illiquidity (Illiq), idiosyncratic volatility (IVOL). We complement this list by a lottery index (LOTT), which is designed to capture the attention of retail investors. To calculate LOTT, we follow Han and Kumar (2013) and also include the maximum daily return over the past one month as an additional input, following Bali et al. (2011).

Low institutional ownership corresponds to a much greater persistence of the mispricing spread between the under- and overvalued stocks. Even after a full year do we find a highly significant spread of 0.91% per month. In contrast, the mispricing spread for the subsample of stocks with a large institutional ownership base is only marginally significant. Small stocks also remain mispriced for a significantly lower period. Interestingly, however, we also find that large stocks can be significantly mispriced, showing that our measure of expected mispricing is not solely valid for small but also large caps.

Table 9: Persistence of Mispricing

The table reports the persistence of mispricing conditional on further firm-characteristics. That is, we sort stocks into decile portfolios based on $\mathbb{E}[MP]$, and, within each $\mathbb{E}[MP]$ decile portfolio further sort stocks into terciles by additional variables. We then keep the portfolio allocation fixed and study the decile-spread of realized mispricing over the next $t + \Delta t$ months for both the low and high tercile of the respective conditioning variable. We include institutional ownership (IO), market capitalization (Size), stock price (PRC), the Amihud (2002) illiquidity measure (Illiq), idiosyncratic volatility (IVOL) and a lottery index (LOTT) designed to capture retail investor's attention. To calculate LOTT we follow Han and Kumar (2013) but include the maximum daily return over the past one month as an additional input (Bali et al., 2011). The table displays the out-of-sample results for the period from January 1993 to December 2022. Data availability restricts IO to end in May 2021.

	IO			Size			PRC		
Δt	Low	High	Δ	Low	High	Δ	Low	High	Δ
1	3.89***	1.63***	−2.26***	3.38***	2.10***	−1.28***	2.96***	2.20***	−0.77
2	2.46***	0.90***	−1.56***	1.89***	1.34***	−0.55	1.61***	1.39***	−0.22
3	1.96***	0.13	−1.83***	1.55***	0.77**	−0.78*	1.83***	0.68**	−1.15**
4	1.96***	0.35	−1.61***	1.31***	0.78**	−0.53	1.51***	0.74*	−0.78*
5	1.06**	0.44*	−0.61	0.85***	0.67***	−0.18	0.86**	0.63**	−0.23
6	2.25***	0.21	−2.05***	1.03***	0.75**	−0.27	0.83**	0.76**	−0.07
7	1.36***	−0.01	−1.37***	0.55**	0.55**	0.01	0.95***	0.53*	−0.42
8	0.95**	0.10	−0.85**	0.84***	0.28	−0.56	0.76**	0.24	−0.52
9	1.24***	−0.01	−1.25***	0.36	0.30	−0.07	1.14***	0.33	−0.81**
10	0.62*	−0.00	−0.62*	0.71***	0.37*	−0.33	0.82**	0.29	−0.53
11	0.92**	0.12	−0.80**	0.59***	0.30	−0.29	0.41	0.38*	−0.03
12	0.91***	−0.70*	−1.61***	0.86***	−0.32	−1.18**	1.11***	−0.47	−1.58***
	Illiq			IVOL			LOTT		
Δt	Low	High	Δ	Low	High	Δ	Low	High	Δ
1	2.06***	2.49***	0.43	2.14***	3.43***	1.29**	2.15***	3.21***	1.06**
2	1.24***	1.57***	0.33	0.83**	1.72***	0.89	0.99***	1.67***	0.67**
3	0.72**	1.20***	0.47	0.99***	1.32***	0.33	0.48	1.62***	1.13**
4	0.76**	1.10***	0.34	0.58*	1.23***	0.65	0.72**	1.24***	0.52
5	0.68***	0.86***	0.18	0.26	1.18**	0.91	0.71***	0.89**	0.18
6	0.71**	0.81***	0.10	0.83***	1.03***	0.20	0.76***	0.98***	0.22
7	0.51*	0.73***	0.23	0.28	0.25	−0.03	0.53*	0.52*	−0.01
8	0.26	0.52**	0.25	0.54**	0.23	−0.31	0.48*	0.42	−0.06
9	0.29	0.86***	0.56*	0.55***	0.95**	0.40	0.12	0.90***	0.78**
10	0.31	0.50**	0.19	0.52**	0.93*	0.41	0.36	0.46	0.10
11	0.26	0.23	−0.03	0.23	0.10	−0.14	0.19	0.13	−0.06
12	−0.44	0.49**	0.93**	0.05	−0.89**	−0.94***	−0.53	0.31	0.84**

A lower nominal price, creating the illusion of a "cheap" investment and thus attracting the attention of retail investors, also leads to a significantly larger persistence of realized mispricing. We do, however, find little evidence of a statistical difference in the mispricing persistence between the stocks with a low and high nominal price. Illiquid stocks also

tend to have a more persistent mispricing. We find an interesting pattern for a stock's idiosyncratic volatility, which Stambaugh et al. (2015) show is a proxy for the costs of engaging in arbitrage activity. High IVOL stocks tend to remain mispriced for longer but the *magnitude* of this mispricing is especially elevated. For example, in the first month after forming the portfolios, the mispricing spread for high IVOL stocks stands at 3.43% compared to 2.14% for low IVOL stocks. In month $\Delta t = 2$, we still find a mispricing spread of 1.72% for high and just 0.83% for low IVOL stocks. These results suggest that IVOL is primarily an indicator of the current costs of arbitrage but that it does not necessarily inform us about the persistence of mispricing. Finally, we show that stocks with characteristics that attract the flows of retail investors (measured by LOTT) also show a more pronounced magnitude of mispricing but are not necessarily mispriced for longer.

4.5. Anomalies & Mispricing

For mispricing to resolve, it must have built up before. We therefore follow van Binsbergen et al. (2023) and classify cross-sectional asset pricing anomalies into those exacerbating mispricing and those that resolve it. As an example, they find that the momentum and profitability anomalies further dislocate prices of stocks from their fair values. To measure these dislocations, van Binsbergen et al. (2023) assume that the SDF is exponentially affine in a set of factors, and consider the CAPM, as well as the Fama and French (1993) model and the Fama and French (2015) model augmented by momentum. Our result in Figure 2 shows that IPCA outperforms these models and better reflects the systematic variation across stock returns. Furthermore, as we discuss in the introduction, our definition of mispricing aligns with the objectives of an arbitrageur by providing a measure of the *timing* of when the mispricing is expected to resolve.

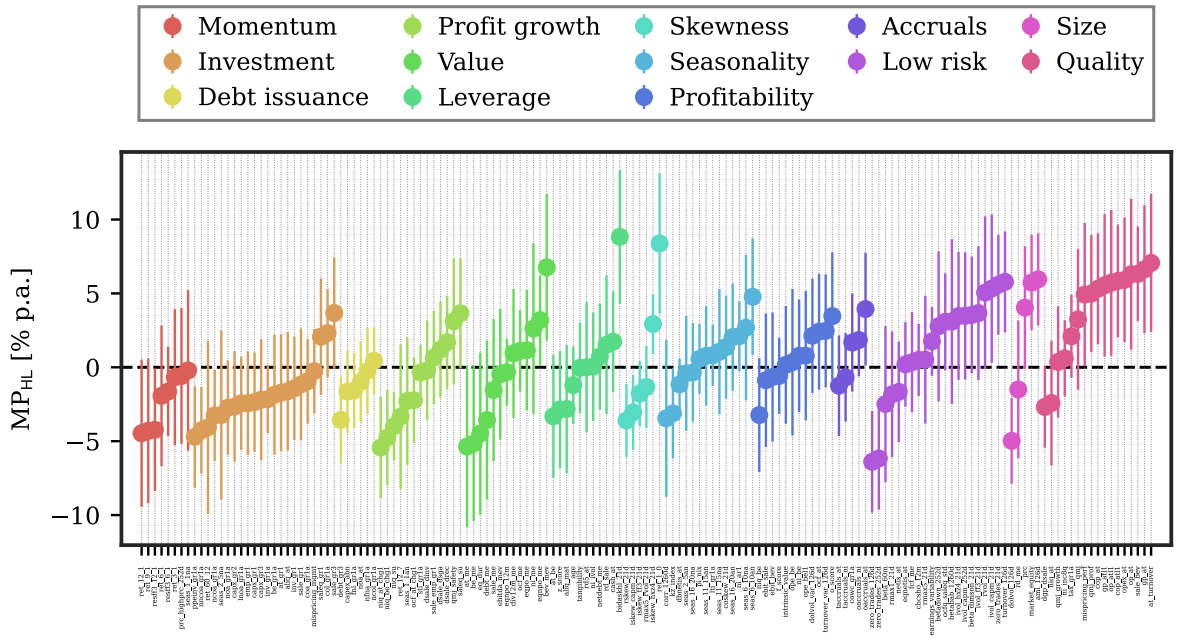


Fig. 5. Mispricing across Characteristic-sorted Portfolios

The figure shows the average high-minus-low spread in realized mispiricing for quintile portfolios individually sorted on 153 different firm-level characteristics provided by Jensen et al. (2023). Within each portfolio, we value-weight stocks with their market capitalization. We sort stocks in a way, such that the unconditional return spread of the high-low portfolio is positive and report the return and the direction of the sorting in Appendix B. Error bars denote the 95% confidence interval from HAC standard errors with 12 lags. We further differentiate characteristics in 13 groups as reported by Jensen et al. (2023). The table displays the out-of-sample results for the period from January 1993 to December 2022.

Our firm-level measure of expected mispricing allows us to revisit this evidence for an extended set of characteristics and a different definition of mispricing. For this, we record the average high-minus-low spread in realized mispricing for quintile portfolios sorted individually on the 153 firm-level characteristics provided by Jensen et al. (2023). We value-weight stocks within each portfolio and sort stocks by each characteristic, such that the unconditional high-minus-low return spread is positive. The sorting direction and average returns are provided in Appendix B. Furthermore, we follow Jensen et al. (2023) and sort the characteristics in 13 groups, which combine characteristics with a similar informational content about the target firm. Examples for these groups are "momentum", "value", or "profitability".

The results are shown in Figure 5. A negative reading indicates that trading on the return anomaly further dislocates prices from their fair values, i.e., that the anomaly proposes buying overvalued and selling undervalued stocks. We find that most momentum-based characteristics further dislocate prices, exacerbating firm-level mispricing, which is consistent with van Binsbergen et al. (2023). However, we find that the evidence is only marginally significant for some momentum proxies and turns insignificant for others. We find that investment-based and some characteristics about the debt issuance of a firm also further dislocate prices. While some profit growth characteristics further dislocate prices, as shown by van Binsbergen et al. (2023), some others, like `qmj_safety` help align market prices with fair values. While most value proxies produce an inconclusive figure of whether they exacerbate or resolve mispricing, two significantly resolve it (`ni_me` and `eqnpo_me`). Trading on the stocks' implied bid-ask spread (`bidaskhl_21d`, group "Leverage") and short-term reversal (`ret_1_0`, group "Skewness") substantially help in realigning market prices, whereas trading on past skewness – a rough measure for the attention of retail investors – instead exacerbates it (`rskew_21d`). Characteristics of a stock's return seasonality and accruals tend to have little impact on the overall mispricing of the traded stocks. Instead, "Low risk" anomalies have the potential to further dislocate *or* align market prices with fair values. Interestingly, trading on the frequency with which zero volume trading days occur over the past month resolves mispricing, whereas trading on the same characteristic but measured over longer periods exacerbates it. Most size-based and quality-oriented characteristics align market prices with fair values and thereby in part resolve the mispricing of the traded stocks.

5. Comparison with Established Measures

This section compares our measure of expected mispricing $\mathbb{E}[MP]$ to the mispricing score by Stambaugh et al. (2015) (SYY) and the anomaly score "Net" by Engelberg et al. (2018), as two established measures of the stock-level mispricing.

Both measures are based on the heuristic that some characteristics inform us about return differences unrelated to risk. For example, SYY is a composite score of the cross-sectional ranks of 11 firm characteristics. However, many of the characteristics have multiple explanations, some of which fall into the "risk-based" category, others in the category of "behavioral biases". An example for this is momentum, for which Kelly, Moskowitz, and Pruitt (2021) propose a risk-based explanation, whereas Barberis, Shleifer, and Vishny (1998) and Hong and Stein (1999) favor an explanation based on the misreaction to news. "Net" generalizes the idea of SYY by considering a vast array of characteristic-based return anomalies, with no pre-selection as to which characteristics may also represent information about systematic risk. For example, Kelly et al. (2019) show that many of the firm characteristics used by Engelberg et al. (2018) in the construction of Net, are useful to adjust stock returns for common sources of risk. As a result, it is unclear to which degree SYY and Net capture mispricing. $\mathbb{E}[MP]$ provides a statistically robust definition of expected mispricing, as the expected deviation from the fair rate of return implied by a statistically-optimal factor model. This section shows that $\mathbb{E}[MP]$ improves upon earlier measures of expected mispricing.

As a first step, in Panel A of Table 10, we compare the value-weighted realized average return spreads of the high-minus-low decile portfolio when stocks are sorted by $\mathbb{E}[MP]$, SYY, or Net. We also show the resulting Sharpe ratios and their statistical significance. $\mathbb{E}[MP]$ generates a high-minus-low average return spread of 23% per year at a Sharpe ratio of 0.89. This performance is roughly equally distributed among the long and the short legs

Table 10: Decile Spreads for Portfolios sorted on Expected Mispricing

The table shows spreads in both realized returns (Panel A) and realized mispricing (Panel B) for decile portfolios sorted on different measures of expected mispricing. We contrast our measure of expected mispricing ($\mathbb{E}$[MP]) against the Stambaugh et al. (2015) mispricing score (SYY) and the Engelberg et al. (2018) anomaly score (Net). We calculate HAC robust standard errors using 12 lags and follow Lo (2002) in calculating the statistics of both the Sharpe and the information ratio. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	Panel A: HML Spreads in Realized Returns					
	$\mathbb{E}$[MP]		SYY		Net	
	r	SR	r	SR	r	SR
Low	-0.103	-0.282	0.112^{***}	0.825^{***}	-0.077	-0.242
High	0.127^{***}	0.547^{*}	-0.023	-0.092	0.119^{***}	0.900^{***}
HL	0.230^{***}	0.890^{***}	-0.135^{***}	-0.764^{***}	0.195^{***}	0.742^{***}

	Panel B: HML Spreads in Realized Mispricing					
	$\mathbb{E}$[MP]		SYY		Net	
	MP	IR	MP	IR	MP	IR
Low	-0.134^{***}	-1.043^{***}	0.013	0.154	0.017	0.157
High	0.121^{***}	1.332^{***}	-0.005	-0.061	-0.004	-0.045
HL	0.255^{***}	1.640^{***}	-0.019	-0.178	-0.021	-0.164

of the arbitrage portfolio. The realized return spreads of SYY and Net are both lower in absolute magnitude, at -13.5% and 19.5%, respectively, with absolute Sharpe ratios north of 0.75. Panel B repeats this exercise but shows the average spreads in *realized mispricing* when sorting by either of the three measures. We also show the information ratio and their statistical significance. Strikingly, $\mathbb{E}$[MP] generates a spread in stocks' realized mispricing of 25.5% per year at an information ratio of 1.64, both of which are highly statistically significant. These magnitudes stand in contrast to the spreads in mispricing uncovered by SYY and Net. Both are small economically and statistically insignificant, showcasing that the two proxies uncover "mispricing" opportunities fully explained by the IPCA factor model. These findings have important practical implications; e.g., an arbitrageur is interested in limiting her exposure to systematic risk and from the three candidate mispricing proxies, only $\mathbb{E}$[MP] allows for that.[6]

[6]Both SYY and Net are unable to uncover economically and statistically meaningful spreads in the

Spanning Regressions. Our measure of mispricing MP is constructed to be unexplained by an optimal latent factor model. We nonetheless investigate if common observable factor models can explain the realized excess returns and the realized mispricing component of the spread portfolios of the three mispricing measures. For that, we perform factor spanning regressions for both the realized mispricing component MP and excess return r of the value-weighted high-minus-low decile portfolio sorted by $\mathbb{E}$[MP], SYY, or Net:

$$\mathrm{MP}^{\mathrm{HL}}_{t+1} = \gamma^{Model} + \delta^{Model} F^{Model}_{t+1} + \eta^{Model}_{t+1} \tag{10}$$

$$r^{\mathrm{HL}}_{t+1} = \alpha^{Model} + \beta^{Model} F^{Model}_{t+1} + \nu^{Model}_{t+1} \tag{11}$$

We are mainly interested in the regressions' constants, which capture the return component unexplained by the exposure to the factors contained in F^{Model}. Here, α^{Model} is the unexplained component of the total excess return and γ^{Model} denotes the unexplained realized mispricing. As candidate models, we include the Fama and French (2015) five-factor model augmented with momentum (Carhart, 1997) (FF6), the Hou, Xue, and Zhang (2015) q-factor model, the Daniel, Hirshleifer, and Sun (2020) behavioral factor model (DHS), the Stambaugh and Yuan (2017) mispricing factor model (SY17), and the six latent factors obtained from IPCA. Finally, we consider a "combined" factor model of all *observable* factors.

We present the results of this analysis in Table 11. Panels A and B report values for γ^{Model} and α^{Model}, respectively. For the decile spread portfolio sorted on $\mathbb{E}$[MP], we find a positive and highly significant γ for all of the factor models considered. For example, the

realized mispricing, which follows our preferred definition based on the IPCA-implied expected return benchmark. Both measures do, however, generate substantial return spreads. In fact, the return spread of Net rivals that of $\mathbb{E}$[MP] at 19.5% per year, showcasing that the anomaly score is able to differentiate high- from low-expected-return stocks. Our results suggest that this return spread is to a large part capturing systematic (risk) variation, and to a lesser degree firm-level mispricing.

Table 11: Spanning Regressions

The table shows spanning regression for both realized mispricing MP (Panel A) and total excess returns r (Panel B) of the decile-spread portfolio based on different mispricing measures. We distinguish our measure of expected mispricing ($\mathbb{E}$[MP]), the Stambaugh et al. (2015) mispricing score (SYY) and the anomaly measure Net of Engelberg et al. (2018). The table reports regression constants α^{Model} and γ^{Model} as outlined in Eq. (11). As candidate factor models we include the Fama and French (2015) five-factor model augmented with momentum(Carhart, 1997) (FF6), the Hou et al. (2015) q-factor model, the Daniel et al. (2020) behavioral factor model (DHS), as well as the Stambaugh and Yuan (2017) mispricing factor model (SY17). We also include a "combined" model of all observable factors. The table displays the out-of-sample results for the period from January 1993 to December 2022. The data for the DHS factors ends in November 2018. Similarly, for SY17 data is available until November 2016.

	Panel A: MP					
	FF6	$q-$Factor	DHS	SY17	IPCA	combined
$\mathbb{E}$[MP]	2.01	1.95	1.77	2.04	1.11	1.82
	[5.70]	[5.40]	[5.07]	[6.04]	[2.74]	[5.86]
SYY	0.27	0.26	0.19	0.47	−0.17	0.46
	[2.13]	[1.64]	[1.21]	[3.10]	[−0.79]	[2.81]
Net	−0.66	−0.77	−0.82	−1.03	−0.40	−1.23
	[−3.29]	[−3.37]	[−4.26]	[−5.76]	[−0.97]	[−6.27]
	Panel B: Total Return					
	FF6	$q-$Factor	DHS	SY17	IPCA	combined
$\mathbb{E}$[MP]	1.50	1.50	1.04	0.93	1.44	0.82
	[4.52]	[3.50]	[3.03]	[2.52]	[2.49]	[2.72]
SYY	−0.67	−0.49	−0.72	−0.16	0.06	−0.12
	[−4.50]	[−2.42]	[−3.38]	[−1.40]	[0.15]	[−1.01]
Net	0.94	0.65	0.77	0.57	−1.05	0.29
	[4.69]	[2.67]	[2.87]	[2.32]	[−1.47]	[1.52]

γ relative to the behavioral factors of Daniel et al. (2020) amounts to 1.77% per month, which is close to the unconditional mispricing spread of 2.13% documented in Table 10. Looking at the unexplained return component in Panel B (α), we likewise find significant values for the $\mathbb{E}$[MP]-sorted spread portfolio for all of the candidate factor models.

Turning to the decile spread portfolio sorted on SYY, we find a significantly positive γ for the FF6, the SY15, and the combined factor models. Remember that a larger reading of SYY corresponds to a relative overvaluation of the respective stock. Following our definition of mispricing, we should thus observe a *negative* mispricing return component in the next month. So even though γ is significant for some factor models, it has the wrong

sign. Looking at the unexplained return component α, we find the correct negative sign. As expected, the SY17 factors, which are constructed from the same set of characteristics as the SYY mispricing score, can explain the resulting spread portfolio returns. We also find that the six latent IPCA factors and the combined model of observable factors also render the α insignificant.

Finally, we investigate the decile spread returns for portfolios sorted on Net. The γ is negative throughout and all cases but IPCA significantly so. However, a higher Net corresponds to a relative *undervaluation*, such that we would expect a *positive* mispricing return component over the next month. We find that Net's α is significantly positive for most models but that both IPCA and the combined observable factor model can explain its return patterns.

In sum, these results show that only $\mathbb{E}[\text{MP}]$ generates mispricing and excess return spreads that (i) are unexplained by common (observable) factors, and (ii) produce the right sign (or accurate cross-sectional relationship). This is suggestive evidence that both SYY and Net capture not only the mispricing component of returns, but also compensation for systematic risk.

Cross-sectional Regressions. We provide supporting evidence for this presumption through stock-level cross-sectional regressions. Each month, we regress the realized mispricing component in month $t+1$ on a constant, $\mathbb{E}[\text{MP}]$, various characteristic-based controls, as well as SYY and Net. For ease of comparison, we standardize all explanatory variables.

The results are shown in Table 12 and confirm our earlier findings from the factor spanning regressions. Realized and expected mispricing ($\mathbb{E}[\text{MP}]$) are tightly connected: a 1% larger $\mathbb{E}[\text{MP}]$ on average translates to a realized mispricing over the next month of 1.3%, irrespective of the regression setup. This effect is highly statistically significant

Table 12: Cross-sectional Predictive Regressions of Realized Mispricing

The table shows results of cross-sectional predictive regressions of realized mispricing (MP) on our measure of expected mispricing $\mathbb{E}[MP]$ including controls. As controls we include the stocks beta (β), the logarithm of its market capitalization (ln(Size)), book-to-market (B2M), momentum (Mom), Investment (Inv), profitability (Prof), Amihud (2002) illiquidity (Illiq), idiosyncratic volatility (Ang et al. (2006), IVOL) and short term reversal (Rev). Additionally, we control for the Stambaugh et al. (2015) mispricing and the Engelberg et al. (2018) anomaly score. All predictors, except for our measure of expected mispricing ($\mathbb{E}[MP]$) are standardized to have unit standard deviation. We calculate HAC robust standard errors using 12 lag. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	(1)	(2)	(3)	(4)	(5)	(6)
const.	0.001	0.004	0.004	0.005	0.005	0.001
	[6.39]	[2.81]	[2.68]	[3.01]	[2.65]	[0.26]
$\mathbb{E}[MP]$	1.309	1.295	1.310	1.319	1.311	1.314
	[10.74]	[11.78]	[11.85]	[11.77]	[11.57]	[6.62]
β		0.000	0.000	−0.000	−0.000	−0.001
		[0.39]	[0.37]	[−0.33]	[−0.39]	[−1.89]
ln(Size)		−0.001	−0.001	−0.001	−0.001	0.001
		[−2.75]	[−2.50]	[−2.47]	[−1.75]	[1.36]
B2M		−0.000	−0.000	−0.000	−0.001	−0.001
		[−0.13]	[−0.48]	[−0.03]	[−1.18]	[−1.01]
Mom			−0.001	−0.001	−0.001	−0.002
			[−2.28]	[−1.96]	[−2.17]	[−3.15]
Inv				−0.002	−0.002	0.002
				[−0.87]	[−0.76]	[1.90]
Prof				0.001	0.001	−0.004
				[0.84]	[0.66]	[−0.53]
Illiq					0.003	−0.005
					[4.53]	[−1.54]
IVOL					−0.000	0.002
					[−0.15]	[1.95]
Rev					0.001	−0.000
					[2.43]	[−0.51]
SYY						−0.000
						[−1.00]
Net						−0.001
						[−2.05]

using Newey and West (1987) standard errors with 12 lags. In regression (2), we include the characteristics of the Fama and French (1993) three-factor model, all of which are cross-sectionally standardized to ease the interpretation. While the log market capitalization turns out to be significantly connected to future mispricing, its inclusion has no impact on the significance and economic magnitude of $\mathbb{E}[MP]$. The same applies to an

addition of stock momentum, the consideration of the characteristics of the Fama and French (2015) five-factor model, or the inclusion of Amihud (2002)'s illiquidity measure, idiosyncratic volatility (IVOL) or short-term reversal (Rev). Finally, we add SYY and Net as alternative mispricing proxies in regression (6). While Net's impact is significant, its coefficient shows the wrong sign; i.e., stocks that Net expects to be overvalued tend to remain overvalued in the next period.

We further test the ability of our expected mispricing measure to predict future realized returns instead of mispricing and report the results in Appendix C. $\mathbb{E}$[MP] remains a significant predictor of future returns in all specifications; that is, its predictive power does not change controlling for either SYY or Net in the multivariate regressions. Reversely, only Net predicts future returns in the specification with all control variables. Given that Table 12 proved Net to have no explanatory power for realized mispricing, it puts further emphasis on the idea that some of Net's ingredients capture considerable factor exposure.

Independent Double Sorts. Finally, we assess the ability of $\mathbb{E}$[MP], and SYY or Net to uncover substantial spreads in next period's realized mispricing, controlling for SYY and Net. For this, we perform independent quintile sorts by $\mathbb{E}$[MP] and SYY in Table 13. The spread between under- and overvalued stocks according to $\mathbb{E}$[MP] is large and highly significant within each SYY quintile. In fact, only for the third $\mathbb{E}$[MP] portfolio do we find a marginally significant mispricing spread with the correct sign for over- vs. undervalued stocks according to SYY. In the upper panel of Table 13, we perform a simple quintile sort by SYY to show that SYY alone is unable to uncover a meaningful spread in the future realized mispricing of stocks.

We perform the same exercise using an independent double sort on $\mathbb{E}$[MP] and Net in Appendix D. The results are roughly comparable to those shown for SYY above. In simple quintile sorts, Net is unable to uncover a meaningful spread in realized mispricing.

Table 13: Expected Mispricing and the Stambaugh et al. (2015) Mispricing Score.

In the upper panel we document the average realized mispricing in month $t+1$ for stocks sorted by the Stambaugh et al. (2015) mispricing score (SYY) measured in month t into quintile portfolios. The lower panel shows the average one-month-ahead realized mispricing for portfolios of stocks sorted independently into quintiles by the Stambaugh et al. (2015) mispricing score (SYY) and our expected mispricing $\mathbb{E}$[MP]. Within each portfolio, we weight stocks by their market capitalization. We calculate t-statistics using HAC robust standard errors. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	SYY					
	Under	2	3	4	Over	HL
MP	0.11	0.12	0.17	0.30^{**}	-0.06	-0.17

	Expected Mispricing					
SYY	Overvalued	2	3	4	Undervalued	HL
Under	-0.27	0.04	0.23	0.28	0.60^{***}	0.88^{***}
2	-0.40^{**}	-0.10	0.04	0.38^{**}	0.66^{***}	1.06^{***}
3	-0.24	0.01	-0.05	0.30^{*}	0.82^{***}	1.06^{***}
4	-0.06	-0.05	-0.04	0.43^{***}	0.97^{***}	1.03^{***}
Over	-0.56^{**}	-0.17	-0.11	0.13	0.34^{**}	0.90^{***}
HL	-0.29	-0.22	-0.34^{*}	-0.15	-0.27	0.02

And even if we control for Net through independent double sorts, we continue to find an economically sizable and highly significant spread in realized mispricing returns between under- and overvalued stocks as identified by $\mathbb{E}$[MP], which ranges from 0.85% to 1.76% per month. Only among overvalued stocks ($\mathbb{E}$[MP]), do we find a significantly positive mispricing spread between the most under- and the most overvalued stocks according to Net.

6. Economy-wide Implications

As a final piece of evidence on the implications of our measure of firm-level *expected mispricing*, we assess the aggregate prevalence of mispricing in the U.S. equity market. We start by recording the relative share of the overall U.S. equity market capitalization

that we currently expect to be overvalued:

$$\text{Share of overvaluation}_t = \frac{\sum_i \mathbb{1}_{\mathbb{E}_t[\text{MP}_{t+1,i}]<0} \times \text{MC}_{t,i}}{\sum_i \text{MC}_{t,i}}. \tag{12}$$

$\text{MC}_{t,i}$ denotes the market capitalization of stock i at time t and $\mathbb{1}_{\mathbb{E}_t[\text{MP}_{t+1,i}]<0}$ is an indicator function that equals 1 if a stock is expected to be overvalued and zero otherwise, as measured by $\mathbb{E}[\text{MP}]$.

Figure 6 shows the time-series of the share of overvaluation. We document several interesting facts. For one, the degree of market-wide overvaluation fluctuates significantly over time. During the 90s, we find that for every dollar invested in the U.S. stock market, between 50 and 80 cents are invested in overvalued stocks. With the burst of the dotcom bubble we find that the degree of market-wide overvaluation drops significantly down to roughly 40% in 2002. We see a similar albeit less pronounced phenomenon in the run-up to the financial crisis: valuations first increase relative to our IPCA benchmark and then come down rapidly as the crisis unfolds in 2008, with an overvaluation share at roughly 40%. For the period after the financial crisis, we find that valuations have mostly remained below 50%. Interestingly, despite lots of talk and media coverage about a new

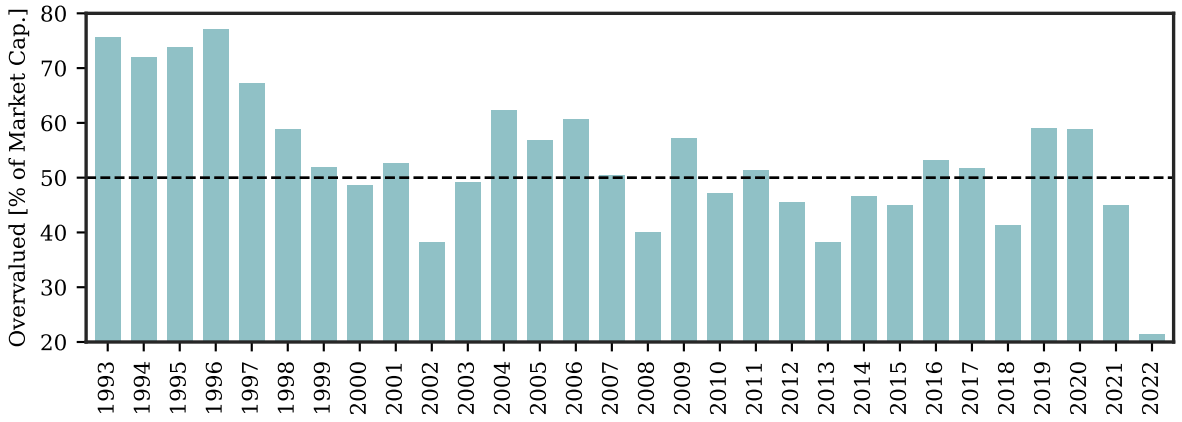


Fig. 6. Fraction of Overvalued Market Capitalization

The figure shows the relative share of the overall U.S. market capitalization that is currently expected to be overvalued (cf. Eq. (12)).

bubble forming, we find that today's share of overvaluation is still substantially lower than during the 90s, with 2022 marking the lowest recorded value of just above 20%.

The results in Figure 6 suggest that the share of overvalued market capitalization co-moves with the business cycle. We formalize this intuition by regressing it on several macroeconomic indicators. We include the intermediary capital ratio of He et al. (2017) as a measure of intermediary health, the market-wide book-to-market ratio, net share issuance, and the treasury bill rate, all of which are taken from Welch and Goyal (2008). We also include the Cboe Volatility Index and the Chicago Fed National Activity Diffusion Index. The regression results are shown in Table 14.

Specification (1) regresses the share of overvaluation on the intermediary capital ratio. If intermediaries are sufficiently capitalized, we expect them to lend more willingly, which in turn amplifies the degree of overvaluation in the economy. We find that a one standard

Table 14: Share of Overvalued Market Capitalization & Macroeconomic Indicators

The table shows the output of regressing the relative share of the overall U.S. market capitalization that is currently expected to be overvalued on several macroeconomic indicators. We include the intermediary capital ratio (ICR) of He et al. (2017), the market-wide book-to-market ratio (b/m), net share issuance ($ntis$), and treasury bill rate (tbl). As further business cycle indicators we include the CBOE Volatility Index VIX as well as the Chicago Fed National Activity Diffusion Index (CFNAI). We standardize both independent and dependent variables and calculate t-statistics using HAC robust standard errors with 6 lags. The table displays the out-of-sample results for the period from January 1993 to December 2022.

	Share of Overvaluation					
	(1)	(2)	(3)	(4)	(5)	(6)
ICR_t	0.24					
	[3.01]					
b/m_t		−0.03				
		[−0.36]				
$ntis_t$			0.35			
			[3.53]			
tbl_t				0.42		
				[4.49]		
VIX_t					−0.36	
					[−3.95]	
CFNAI_t						0.26
						[3.06]

deviation increase in *ICR* leads to an increase in the fraction of overvaluation of 0.24 standard deviations, confirming our intuition. Specification (2) includes the book-to-market ratio of the Dow Jones Industrial Average (b/m). A higher book-to-market ratio indicates a lower valuation of the aggregate stock market, relative to the market's book value. This should correspond to a lower degree of overvaluation. While we find a negative coefficient in line with this notion, this dependence does not turn out to be statistically significant. We do, however, find a significant relationship for the market-wide net equity share issuance (*ntis*) in specification (3). A higher net share issuance should coincide with a higher degree of overvaluation, should managers strategically exploit their elevated share prices (Pontiff and Woodgate, 2008; Greenwood and Hanson, 2012; McLean et al., 2020). Additionally, we include the current Treasury bill rate *tbl* in specification (4), as a rough proxy for overall lending costs, and a benchmark for the costs of short-selling. If short-selling is more costly, we expect more stocks to be overvalued. We also confirm this idea in the data with a one standard deviation increase in *tbl* elevating the overvaluation share by 0.42 standard deviations. Higher levels of overall uncertainty, which we measure using the CBOE volatility index (VIX) decrease the share of overvaluation, consistent with the idea that times of lower investor sentiment and especially crises resolve mispricing, shown in specification (5). Finally, we proxy for the business cycle using the Chicago Fed National Activity Diffusion Index (CFNAI).[7] Consistent with the idea that growth expectations might be overly optimistic in good economic states, we find a significantly positive coefficient for the CFNAI index in specification (6).

[7] Brave, Lichtenstein, et al. (2012) show that the CFNAI Diffusion index is a more timely predictor of changes in economic conditions than the traditional CFNAI index.

7. Conclusion

We introduce a novel measure of firm-level *expected mispricing*, $\mathbb{E}[\text{MP}]$, which is grounded in economic theory and closely follows the real-time investment execution process of arbitrageurs. For this, we first propose a general definition of realized mispricing as the deviation from the fair rate of excess returns implied by the stochastic discount factor. To approximate the stochastic discount factor, we use a statistical factor model that optimally extracts common sources of return variation from the cross-section of stock returns. In a second step, we employ modern machine learning predictors to find a mapping between a large set of observable firm characteristics and the expected mispricing component of stock returns.

We find that our setup adequately captures the dynamics of firm-level mispricing and generates a value-weighted return spread in realized mispricing of 2.13% per month, relative to the optimal factor model between stocks identified as undervalued vs. stocks identified as overvalued. The spreads in average raw returns and risk-adjusted returns (FF6 alphas) on $\mathbb{E}[\text{MP}]$-sorted portfolios are comparably large at 1.92% and 1.50% per month, respectively. Small and illiquid stocks with high investor disagreement, low profitability and low nominal share prices are more likely to be overvalued. Furthermore, the prevalence and the degree of mispricing is related to various proxies for limits-to-arbitrage and informational frictions. We show that an overvaluation is amplified after periods of high investor sentiment (Stambaugh et al., 2012), and that earnings announcements play a pivotal role in correcting the mispricing of individual stocks (Engelberg et al., 2018). Finally, we investigate the persistence of mispricing and – contrary to popular belief – find that the magnitude of an over- and an undervaluation is corrected at a similar pace. Once we condition on the prevalence of short-sale constraints, we find that it takes significantly longer to correct an overvaluation of more constrained stocks.

We compare $\mathbb{E}$[MP] with the mispricing score of Stambaugh et al. (2015) and the anomaly score by Engelberg et al. (2018) and find in factor spanning regressions that only $\mathbb{E}$[MP] generates return and mispricing spreads that are unexplained by leading (observable) factor models and of the expected sign. We provide supporting evidence for this in firm-level cross-sectional regressions and bivariate portfolio sorts on $\mathbb{E}$[MP], and SYY or Net.

Finally, we briefly touch on the implications of $\mathbb{E}$[MP] for the aggregate U.S. economy. We find that the share of overvalued market capitalization fluctuates significantly over time, is elevated ahead of major market selloffs and diminishes afterwards. The share of overvalued market capitalization is high whenever lending conditions are good, managers perceive valuations to be high, and economic conditions are expected to improve.

References

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

Ang, A., Hodrick, R. J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. The Journal of Finance 61, 259–299.

Atilgan, Y., Bali, T. G., Demirtas, K. O., Gunaydin, A. D., 2020. Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns. Journal of Financial Economics 135, 725–753.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. The Journal of Finance 61, 1645–1680.

Bali, T. G., Beckmeyer, H., Moerke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. Review of Financial Studies 36, 3548–3602.

Bali, T. G., Cakici, N., Whitelaw, R. F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. Journal of Financial Economics 99, 427–446.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Bartram, S. M., Grinblatt, M., 2018. Agnostic fundamental analysis works. Journal of Financial Economics 128, 125–147.

Birru, J., Wang, B., 2016. Nominal price illusion. Journal of Financial Economics 119, 578–598.

Brave, S., Lichtenstein, M., et al., 2012. A different way to review the Chicago Fed National Activity Index. Chicago Fed Letter May, Federal Reserve Bank of Chicago.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

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

Daniel, K., Hirshleifer, D., Sun, L., 2020. Short-and long-horizon behavioral factors. The Review of Financial Studies 33, 1673–1736.

Diether, K. B., Malloy, C. J., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. The Journal of Finance 57, 2113–2141.

Engelberg, J., McLean, R. D., Pontiff, J., 2018. Anomalies and news. The Journal of Finance 73, 1971–2001.

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

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

Friedman, J. H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics pp. 1189–1232.

Greenwood, R., Hanson, S. G., 2012. Share issuance and factor timing. The Journal of Finance 67, 761–798.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Han, B., Kumar, A., 2013. Speculative retail trading and asset prices. Journal of Financial and Quantitative Analysis 48, 377–404.

Han, Y., He, A., Rapach, D., Zhou, G., 2023. Cross-sectional expected returns: New fama-macbeth regressions in the era of machine learning. Available at SSRN 3185335 .

He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics 126, 1–35.

Hong, H., Stein, J. C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of finance 54, 2143–2184.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. The Review of Financial Studies 28, 650–705.

Jensen, T. I., Kelly, B. T., Pedersen, L. H., 2023. Is there a replication crisis in finance? Journal of Finance 78.

Kelly, B. T., Moskowitz, T. J., Pruitt, S., 2021. Understanding momentum and reversal. Journal of Financial Economics 140, 726–743.

Kelly, B. T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524.

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting factor models. The Journal of Finance 73, 1183–1223.

Lo, A. W., 2002. The statistics of sharpe ratios. Financial Analysts Journal 58, 36–52.

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

McLean, R. D., Pontiff, J., Reilly, C., 2020. Taking sides on return predictability. Working Paper, Available at SSRN 3637649.

Miller, E. M., 1977. Risk, uncertainty, and divergence of opinion. The Journal of Finance 32, 1151–1168.

Nagel, S., 2005. Short sales, institutional investors and the cross-section of stock returns. Journal of Financial Economics 78, 277–309.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. The Journal of Finance 63, 921–945.

Rapach, D. E., Strauss, J. K., Zhou, G., 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. The Review of Financial Studies 23, 821–862.

Stambaugh, R. F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.

Stambaugh, R. F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. The Journal of Finance 70, 1903–1948.

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

van Binsbergen, J. H., Boons, M., Opp, C. C., Tamoni, A., 2023. Dynamic asset (mis) pricing: Build-up versus resolution anomalies. Journal of Financial Economics 147, 406–431.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21, 1455–1508.

Zhang, X. F., 2006. Information uncertainty and stock returns. Journal of Finance 61, 105–137.

Zhou, G., Zhou, Z. C., 2023. What explains gamestop price surge? evidence from both stock and option tick-by-tick trading. Working paper, Available at SSRN.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology 67, 301–320.

Appendix

A. The Impact of Size on Mispricing Spreads

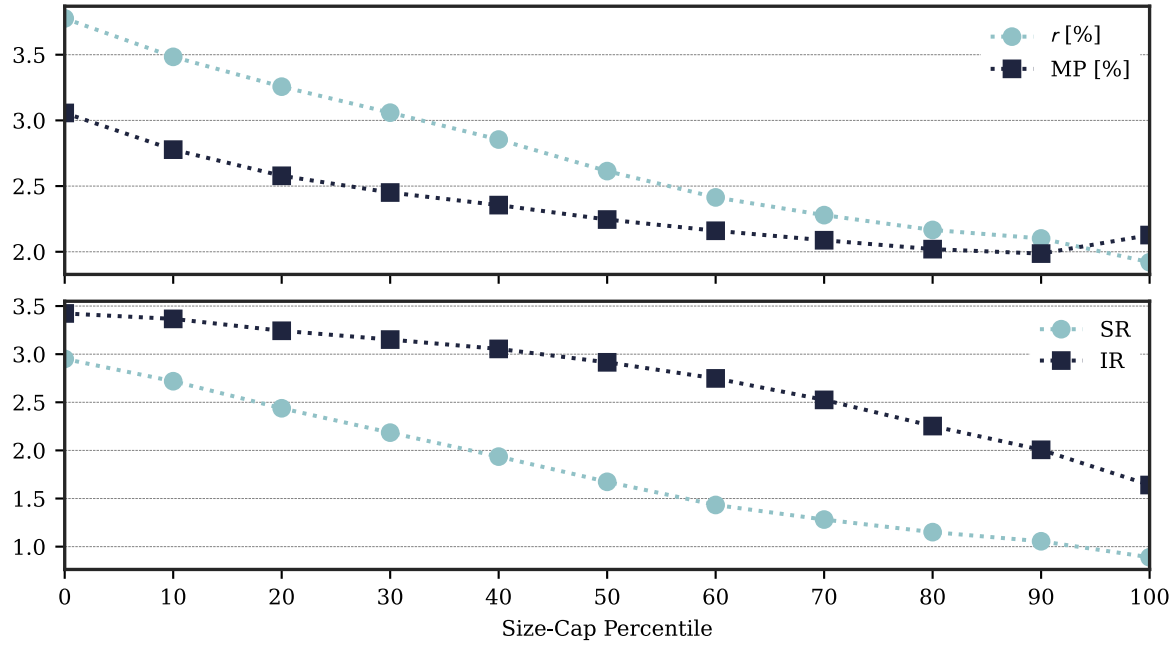


Fig. A.1. Impact of Size

The figure shows performance metrics of the high-low spread portfolio for a decile sort on expected mispricing as a function of a varying size-cap percentile. That is, we cap the weights at the xth percentile of the market capitalization distribution, such that the weight of stocks with higher market capitalization is set to the value of the xth percentile. A size-cap percentile of zero effectively resembles an equal weighting. We show results for the realized return (r), realized mispricing (MP), Sharpe ratio (SR) and Information ratio (IR).

B. Mispricing Across Characteristic-sorted Portfolios

The table shows the time-series average expected mispricing ($\mathbb{E}[MP]$) and realized mispricing (MP) for quintile portfolios sorted on 153 different firm characteristics provided by Jensen et al. (2023). We value-weight stocks with their market capitalization within each portfolio an report values for the lowest (Low) and highest (High) quintile portfolio, as well as the spread between both portfolios (HL). We sort stocks in a way, such that the unconditional return spread of the high-low portfolio (r_{HL}) within our sample is positive. We highlight the direction of the sorting in the column denoted by "Sign". Standard errors are HAC robust using 12 lags.

	$\mathbb{E}[MP]$			MP				
	Low	High	HL	Low	High	HL	Sign	r_{HL}
age	0.96***	−0.66***	−1.61***	2.20	0.99	−1.21	1	1.34
aliq_at	0.91***	−0.49**	−1.40***	3.17	1.55	−1.62	−1	1.64
aliq_mat	−0.09	−0.06	0.04	2.52	−0.30	−2.82	1	1.54
ami_126d	−3.10***	−0.12	2.98***	−4.42***	1.54	5.95***	−1	1.62
at_be	0.56***	−0.47*	−1.04***	1.88	−1.44	−3.32	1	2.16
at_gr1	0.78***	−0.30	−1.07***	2.04	0.33	−1.72	−1	5.86*
at_me	0.51***	−0.83***	−1.34***	3.32	−2.04	−5.36*	1	0.98
at_turnover	−0.80***	0.89***	1.69***	−3.28	3.79**	7.06***	1	3.28
be_gr1a	0.94***	−0.08	−1.02***	4.15*	2.29	−1.86	−1	2.75
be_me	0.42***	−0.46**	−0.89***	2.87	−2.29	−5.15*	1	0.86
beta_60m	−0.53***	0.74***	1.26***	2.20	−0.28	−2.49	1	2.04
beta_dimson_21d	0.94***	0.26**	−0.68***	−0.70	2.84	3.54*	−1	3.20
betabab_1260d	0.70***	−0.92***	−1.62***	−0.24	2.86**	3.10	−1	1.81
betadown_252d	0.95***	−0.62***	−1.57***	0.25	3.03**	2.78	−1	2.02
bev_mev	−0.48***	0.47***	0.95***	−3.76*	3.01	6.76***	−1	0.77
bidaskhl_21d	−4.44***	−0.76***	3.69***	−7.23***	1.60	8.83***	−1	12.51**
capex_abn	0.53***	0.28*	−0.26**	2.42	0.77	−1.65	−1	2.44
capx_gr1	0.85***	0.00	−0.84***	1.83	−0.53	−2.36	−1	6.07**
capx_gr2	1.16***	−0.07	−1.23***	3.67*	1.02	−2.66	−1	5.96**
capx_gr3	1.11***	−0.27	−1.38***	2.16	−0.04	−2.19	−1	4.29
cash_at	−0.10	0.46***	0.56**	1.61	3.34	1.73	1	3.12
chcsho_12m	0.22*	−0.33	−0.55***	1.73	2.23*	0.50	−1	4.69*
coa_gr1a	0.58***	0.09	−0.49***	2.35	1.40	−0.95	−1	4.01**
col_gr1a	−0.15	0.84***	0.99***	1.04	3.35*	2.31	1	1.07
cop_at	−0.65***	0.23	0.88***	−2.11	3.20*	5.31***	1	10.37***
cop_atl1	−0.67***	0.45***	1.12***	−2.33	3.51*	5.84***	1	9.72***
corr_1260d	−0.68***	−0.41**	0.26	3.60	0.15	−3.45	1	2.22
coskew_21d	−0.14	−0.07	0.06	0.28	1.65	1.36	−1	3.01
cowc_gr1a	0.08	0.73***	0.65***	0.77	2.45	1.67	−1	6.31***
dbnetis_at	−0.17	0.31**	0.48***	2.00	0.84	−1.16	−1	0.87
debt_gr3	0.55***	0.28**	−0.27**	3.30*	−0.26	−3.56**	−1	1.10
debt_me	0.61***	−0.65***	−1.26***	1.53	−2.04	−3.57	1	0.12
dgp_dsale	0.11	0.08	−0.04	1.50	−1.18	−2.68*	1	0.45
div12m_me	−0.82***	0.77***	1.59***	1.74	2.68	0.94	−1	1.45
dolvol_126d	−0.09	−2.72***	−2.63***	1.59	−3.39***	−4.98***	−1	1.12

Continued on next page.

	$\mathbb{E}$[MP]			MP				
	Low	High	HL	Low	High	HL	Sign	r_{HL}
dolvol_var_126d	−1.18***	−0.29	0.89***	−0.90	1.23	2.13	−1	1.37
dsale_dinv	0.04	−0.07	−0.10	0.52	0.31	−0.21	1	2.67
dsale_drec	0.47***	0.24	−0.23***	0.99	2.69*	1.70	−1	1.69
dsale_dsga	0.44***	0.50***	0.05	0.60	1.79	1.19	−1	0.96
earnings_variability	−0.26	−0.13	0.13	−0.31	1.45	1.75	1	1.37
ebit_bev	−0.31	−0.02	0.28	2.85	2.19	−0.65	1	7.88
ebit_sale	−0.46*	−0.33	0.13	1.87	1.00	−0.87	1	7.66
ebitda_mev	−1.20***	0.05	1.26***	1.35	0.87	−0.47	1	10.45*
emp_gr1	0.76***	−0.22	−0.97***	2.25	−0.18	−2.43	−1	3.17
eq_dur	1.26***	−0.09	−1.35***	3.74	−0.74	−4.49	−1	3.42
eqnetis_at	0.63***	−0.28	−0.91**	2.24	2.44	0.19	−1	7.48*
eqnpo_12m	0.29**	−0.45**	−0.74***	2.12	1.79	−0.33	1	4.20
eqnpo_me	−0.20	−0.37	−0.17	−0.38	2.81**	3.19**	1	9.43**
eqpo_me	0.08	−0.29	−0.37	1.20	2.36*	1.16	1	3.40
f_score	0.30*	−0.08	−0.38**	2.22	1.64	−0.58	1	2.15
fcf_me	−1.12***	−0.00	1.12***	−1.35	−1.01	0.34	1	7.42**
fnl_gr1a	0.24	0.11	−0.13	2.17	0.61	−1.57	−1	3.27**
gp_at	−0.94***	0.27	1.21***	−2.86	3.77**	6.63***	1	7.20***
gp_atl1	−0.90***	0.51***	1.41***	−1.85	3.70*	5.55**	1	3.97
intrinsic_value	−0.57***	−0.30	0.26	1.25	1.41	0.17	1	0.90
inv_gr1	0.25*	−0.33	−0.58***	1.09	−0.33	−1.43	−1	3.70*
inv_gr1a	0.58***	−0.07	−0.65***	3.19**	1.04	−2.15*	−1	4.77***
iskew_capm_21d	1.19***	−0.92***	−2.12***	2.98**	−0.07	−3.05**	−1	0.48
iskew_ff3_21d	1.08***	−0.83***	−1.91***	2.45*	0.66	−1.79	−1	0.25
iskew_hxz4_21d	−0.79***	1.00***	1.79***	0.23	3.16*	2.93***	1	1.62
ivol_capm_21d	−4.28***	−1.25***	3.03***	−4.08*	1.24	5.32**	−1	11.80**
ivol_capm_252d	−4.15***	−0.82***	3.33***	−2.57	0.91	3.48	−1	10.59
ivol_ff3_21d	−4.18***	−1.22***	2.96***	−2.21	1.46	3.67	−1	9.67*
ivol_hxz4_21d	−4.22***	−1.20***	3.02***	−2.30	1.17	3.47	−1	9.38*
kz_index	0.99***	−0.64***	−1.62***	3.39*	0.27	−3.11**	−1	0.13
lnoa_gr1a	0.59***	−0.41**	−1.00***	2.42	−0.02	−2.44	−1	4.66**
lti_gr1a	−0.12	−0.51**	−0.38***	0.24	1.05	0.81	−1	1.46
market_equity	−4.00***	−0.11	3.89***	−4.25***	1.49	5.73***	1	1.65
mispricing_mgmt	0.45***	−0.47**	−0.92***	1.36	1.10	−0.26	1	7.55***
mispricing_perf	−1.69***	−0.04	1.64***	−2.64	2.28	4.92**	1	13.82***
ncoa_gr1a	0.64***	−0.26	−0.89***	3.71*	0.45	−3.25**	−1	3.90*
ncol_gr1a	−0.15	−0.17	−0.02	1.43	1.87	0.44	−1	0.92
netdebt_me	−0.52**	−0.51***	0.01	−1.45	−0.72	0.72	−1	3.49
netis_at	0.39**	0.10	−0.29	2.83	1.16	−1.66	−1	6.87
nfna_gr1a	0.24	0.32**	0.08	2.28	2.04	−0.24	1	4.30***
ni_ar1	−0.43**	−0.12	0.31***	0.28	2.39	2.11*	1	1.98
ni_be	−0.03	−0.23	−0.20	1.18	1.96	0.78	1	5.92
ni_inc8q	−0.06	−0.16	−0.10	1.65	2.25	0.60	−1	0.49
ni_ivol	0.55***	−0.73***	−1.27***	0.13	0.18	0.05	−1	1.03
ni_me	−1.93***	−0.27	1.66***	−1.47	1.14	2.61	1	8.64
niq_at	−0.04	−0.07	−0.03	4.91**	2.50	−2.42	1	6.12
niq_at_chg1	1.19***	−0.18	−1.37***	4.52**	−0.91	−5.43***	1	0.42

Continued on next page.

	$\mathbb{E}$[MP]			MP				
	Low	High	HL	Low	High	HL	Sign	r_{HL}
niq_be	0.54***	−0.26	−0.80***	4.78**	1.55	−3.23	1	5.69
niq_be_chg1	1.08***	−0.36**	−1.45***	3.91**	−0.85	−4.76***	1	2.32
niq_su	0.51***	−0.45***	−0.96***	3.61**	−0.38	−3.99***	1	1.61
nncoa_gr1a	0.72***	−0.36*	−1.08***	4.45**	0.19	−4.27***	−1	4.03*
noa_at	0.65***	−0.01	−0.66***	2.08	1.09	−0.99	−1	8.14***
noa_gr1a	0.77***	−0.03	−0.81***	3.24	0.52	−2.73*	−1	6.61***
o_score	−0.99***	−0.13	0.86**	−1.36	2.11	3.46	−1	7.20
oaccruals_at	−0.08	0.81***	0.89***	−1.10	2.84	3.94**	−1	6.38**
oaccruals_ni	−0.20	0.53**	0.73***	−0.71	1.14	1.85	−1	4.62***
ocf_at	−0.93***	0.08	1.02***	0.23	2.64	2.41	1	8.67**
ocf_at_chg1	0.69***	0.49***	−0.20**	2.65*	0.42	−2.23	1	0.59
ocf_me	−1.41***	0.05	1.47***	−1.20	−0.09	1.11	1	8.11***
ocfq_saleq_std	−0.10	−0.53**	−0.43**	−0.38	2.70**	3.08*	−1	2.85
op_at	−1.34***	0.11	1.45***	−3.59	2.71	6.29**	1	13.39***
op_atl1	−1.35***	0.27*	1.62***	−2.71	3.01	5.72**	1	11.90***
ope_be	−0.33	−0.04	0.29	2.10	2.44	0.34	1	9.41*
ope_bel1	−0.43*	0.12	0.55*	1.57	2.36	0.79	1	8.43*
opex_at	−0.49**	0.86***	1.36***	−2.08	3.84**	5.92***	1	3.10
pi_nix	0.18	0.27	0.08	0.92	1.49	0.57	1	1.57
ppeinv_gr1a	1.09***	−0.42**	−1.51***	4.42**	−0.29	−4.71***	−1	4.20*
prc	−4.79***	−0.19	4.61***	−2.32	1.72	4.04*	1	4.17
prc_highprc_252d	−2.27***	0.04	2.30***	1.11	0.53	−0.59	1	8.53
qmj	0.13	−0.23	−0.37**	−0.76	2.48	3.23	1	5.66*
qmj_growth	0.14	0.09	−0.05	1.92	2.29	0.36	1	3.97*
qmj_prof	−0.68***	0.04	0.73***	−2.05	2.94*	4.98**	1	10.45***
qmj_safety	0.16	−0.24*	−0.40***	−1.10	2.01	3.11	1	2.89
rd5_at	0.67***	0.51***	−0.17	4.97***	4.98**	0.00	1	7.31
rd_me	1.02***	−0.61***	−1.63***	3.83**	2.33	−1.50	1	8.98**
rd_sale	0.55**	1.08***	0.53*	3.39	4.91***	1.52	−1	2.10
resff3_12_1	0.42**	−0.27	−0.69***	3.84**	−0.38	−4.23**	1	7.29**
resff3_6_1	0.11	−0.23	−0.33***	2.68	1.04	−1.64	1	4.33**
ret_12_1	−0.44**	0.97***	1.41***	4.44**	−0.03	−4.46*	1	9.68**
ret_12_7	0.57***	0.77***	0.19	4.91***	1.57	−3.34	1	6.06
ret_1_0	−1.39***	2.69***	4.08***	−4.04**	4.34***	8.38***	1	0.30
ret_3_1	−0.52***	0.98***	1.50***	2.96	2.28	−0.68	1	5.03*
ret_60_12	0.47***	−0.02	−0.49**	2.85	−1.21	−4.07	−1	1.69
ret_6_1	−0.80***	1.00***	1.80***	2.23	0.30	−1.93	1	8.35**
ret_9_1	−0.60***	1.08***	1.67***	5.04***	0.74	−4.30*	1	10.25**
rmax1_21d	0.41*	−1.58***	−1.98***	3.00	1.17	−1.83	−1	8.33
rmax5_21d	−0.98***	−1.73***	−0.75**	0.85	1.37	0.52	−1	10.92*
rmax5_rvol_21d	1.69***	−1.76***	−3.46***	2.93**	1.60	−1.33	−1	6.91***
rskew_21d	1.26***	−0.90***	−2.16***	3.77***	0.18	−3.60***	−1	1.47
rvol_21d	−3.66***	−1.42***	2.25***	−3.52	1.53	5.05*	−1	11.85*
sale_bev	−0.50**	0.77***	1.26***	−1.70	4.62***	6.33***	1	8.12***
sale_emp_gr1	0.52***	0.08	−0.44***	0.82	1.48	0.65	−1	1.01
sale_gr1	0.84***	0.01	−0.83***	2.04	0.88	−1.16	−1	2.78
sale_gr3	−0.34*	0.98***	1.33***	−0.33	3.35	3.68*	1	0.05

Continued on next page.

	$\mathbb{E}$[MP]			MP				
	Low	High	HL	Low	High	HL	Sign	r_{HL}
sale_me	0.65^{***}	0.38^{**}	-0.27^{*}	2.91	1.37	−1.54	1	5.72
saleq_gr1	0.43^{**}	0.63^{***}	0.20	1.20	3.28	2.07	1	1.13
saleq_su	−0.29	0.41^{**}	0.70^{***}	−0.24	3.41^{**}	3.65^{*}	−1	0.12
seas_11_15an	−0.11	0.02	0.12^{*}	2.31	3.08^{*}	0.77	1	2.90
seas_11_15na	0.02	0.02	0.00	1.53	2.59^{*}	1.05	−1	2.34
seas_16_20an	−0.30	−0.15	0.15^{**}	0.48	2.52	2.04	1	3.87^{*}
seas_16_20na	0.23	-0.70^{***}	-0.93^{***}	3.24^{*}	2.90^{*}	−0.35	−1	0.25
seas_1_1an	0.58^{***}	0.41^{***}	-0.16^{*}	3.11^{*}	0.82	−2.29	1	2.28
seas_1_1na	−0.05	2.17^{***}	2.22^{***}	3.12	2.90	−0.21	1	8.34^{*}
seas_2_5an	0.20	0.62^{***}	0.42^{***}	2.30	1.89	−0.41	1	1.26
seas_2_5na	0.90^{***}	−0.07	-0.96^{***}	2.39	−0.84	−3.23	−1	2.02
seas_6_10an	−0.11	0.40^{*}	0.51^{***}	−1.68	3.09	4.77^{**}	1	8.26^{***}
seas_6_10na	0.46^{*}	0.41^{**}	−0.05	1.29	3.97^{***}	2.68	−1	6.06^{**}
sti_gr1a	0.10	-0.32^{*}	-0.42^{***}	1.60	1.25	−0.35	−1	0.54
taccruals_at	0.60^{***}	0.26^{**}	-0.33^{***}	2.29	1.05	−1.25	1	0.72
taccruals_ni	0.31^{*}	0.15	−0.16	0.87	0.18	−0.68	1	0.71
tangibility	−0.19	0.01	0.19	1.65	1.61	−0.04	1	0.88
tax_gr1a	0.24^{*}	0.21	−0.04	1.29	3.39^{**}	2.10	−1	0.20
turnover_126d	-1.62^{***}	1.40^{***}	3.02^{***}	−1.31	4.48^{**}	5.78^{***}	1	0.85
turnover_var_126d	-1.26^{***}	−0.25	1.02^{***}	−1.03	1.44	2.46	−1	2.43
z_score	0.44^{***}	-0.62^{***}	-1.06^{***}	2.49	−0.36	−2.86	−1	0.38
zero_trades_126d	1.37^{***}	-1.68^{***}	-3.05^{***}	4.36^{**}	−2.04	-6.40^{***}	1	0.54
zero_trades_21d	-1.73^{***}	1.39^{***}	3.12^{***}	-2.61^{**}	2.98	5.59^{***}	−1	1.13
zero_trades_252d	1.33^{***}	-1.63^{***}	-2.96^{***}	4.42^{**}	−1.75	-6.17^{***}	1	0.76

C. Cross-sectional Regression – Future Realized Returns

The table shows results of cross-sectional predictive regressions of realized returns (r) on our measure of expected mispricing $\mathbb{E}[MP]$ including controls. As controls we include the stocks beta (β), the logarithm of its market capitalization (ln(Size)), book-to-market (B2M), momentum (Mom), Investment (Inv), profitability (Prof), Amihud (2002) illiquidity (Illiq), idiosyncratic volatility (Ang et al. (2006), IVOL) and short term reversal (Rev). Additionally, we control for the Stambaugh et al. (2015) mispricing and the Engelberg et al. (2018) anomaly score. All predictors, except for our measure of expected mispricing ($\mathbb{E}[MP]$) are standardized to have unit standard deviation. We calculate HAC robust standard errors using 12 lags.

	(1)	(2)	(3)	(4)	(5)	(6)
const.	0.009	0.015	0.014	0.015	0.018	0.013
	[2.76]	[3.32]	[3.21]	[3.08]	[4.39]	[4.58]
$\mathbb{E}[MP]$	1.583	1.515	1.493	1.515	1.337	1.014
	[10.61]	[12.95]	[13.05]	[13.21]	[13.00]	[6.48]
β		0.001	0.001	0.000	0.001	0.001
		[0.54]	[0.49]	[0.36]	[0.60]	[0.99]
ln(Size)		−0.002	−0.003	−0.003	−0.003	−0.001
		[−2.24]	[−2.52]	[−2.45]	[−3.39]	[−2.34]
B2M		0.001	0.001	0.001	0.001	0.001
		[0.85]	[1.41]	[1.43]	[0.67]	[0.40]
Mom			0.002	0.002	0.002	0.000
			[2.65]	[2.36]	[2.42]	[0.37]
Inv				−0.004	−0.004	0.001
				[−1.44]	[−1.44]	[1.72]
Prof				0.004	0.003	−0.010
				[2.64]	[2.46]	[−1.03]
Illiq					0.003	−0.006
					[4.72]	[−1.23]
IVOL					−0.003	−0.001
					[−2.86]	[−1.11]
Rev					−0.002	−0.004
					[−2.95]	[−4.38]
SYY						−0.001
						[−1.63]
Net						0.004
						[5.08]

D. Expected Mispricing and the Anomaly Score "Net"

In the upper panel we document the average realized mispricing in month $t+1$ for stocks sorted by the Engelberg et al. (2018) anomaly score (Net) measured in month t into quintile portfolios. The lower panel shows the average one-month-ahead realized mispricing for portfolios of stocks sorted independently into quintiles by the Engelberg et al. (2018) anomaly score (Net) and our expected mispricing $\mathbb{E}$[MP]. Within each portfolio, we weight stocks by their market capitalization. We calculate t-statistics using HAC robust standard errors.

	Net					
	Over	2	3	4	Under	HL
MP	0.14	0.20	0.15	0.11	0.05	−0.10

	Expected Mispricing					
Net	Overvalued	2	3	4	Undervalued	HL
Over	−0.97***	−0.57***	−0.24	0.36	0.79***	1.76***
2	−0.62***	−0.30	−0.04	0.12	0.90***	1.52***
3	−0.64***	−0.09	0.04	0.21**	0.66***	1.30***
4	−0.24	−0.20	0.10	0.23	0.62***	0.85***
Under	−0.38**	−0.16	0.09	0.25	0.71***	1.09***
HL	0.59**	0.41*	0.34	−0.11	−0.08	−0.67*

E. Persistence of Mispricing - Graphical Illustration

We analyze the persistence of mispricing in Section 4.4 and find that the magnitude of an over- and an undervaluation is corrected at a similar pace. At first, this may seem at odds with conventional wisdom in that an overvaluation should be purged from the market more slowly in the presence of costly short-selling (Stambaugh et al., 2015). Here, we extend the analysis for the additional institutional ownership split graphically by looking at the most over- and undervalued portfolio separately, and investigate the role of *binding* short-selling constraints. That is, we investigate the persistence patterns of stocks that are more likely to be short-selling constrained. We follow Nagel (2005) and sort the stocks within each $\mathbb{E}$[MP] decile portfolio into terciles by their institutional ownership share. Since institutional investors commonly facilitate the short-selling of stocks, we expect more persistent overvaluation among stocks with high retail ownership.

The results in Figure E.1 agree with this; i.e., overvalued stocks with more binding short-sale constraints remain overvalued for a longer period of time. Specifically, we find that their overvaluation persists even after 12 months. In contrast, the overpricing of stocks with a heavy presence of institutional investors is corrected in as little as $\Delta t = 2$ months. We find only small differences in the persistence of mispricing among undervalued stocks with a high and a low institutional ownership share.

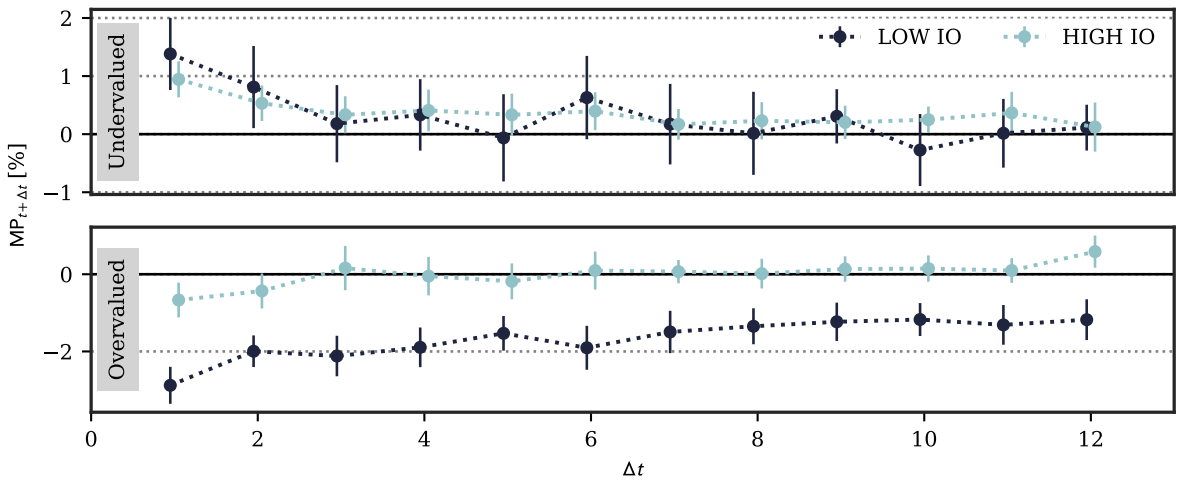


Fig. E.1. Mispricing Persistence – Conditional on Short-Selling Constraints

The figure shows the decay in mispricing over time. In month $t = 0$ we first sort stocks into deciles by their expected mispricing $\mathbb{E}$[MP]. Within each decile, we then sort stocks into terciles by their institutional ownership share and record the future realized mispricing in months $t+1$ to $t+12$. We plot the results for the most undervalued and overvalued stocks and split the results in the highest and lowest institutional ownership tercile for each of the two decile portfolios. Errorbars denote the 5% confidence levels.