The Review of Financial Studies

Machine Forecast Disagreement

Turan G. Bali
Georgetown University, the United States

Bryan T. Kelly
Yale University, AQR Capital Management, and NBER, the United States

Mathis Mörke
ESCP Business School Paris, France

Jamil Rahman
Yale University, the United States

We propose a statistical model of heterogeneous beliefs wherein investors are represented as different machine learning model specifications. Investors form return forecasts from their individual models using common data inputs. We measure disagreement as forecast dispersion across investor-models (MFD). Our measure aligns with analyst forecast disagreement but more powerfully predicts returns. We document a large and robust association between belief disagreement and future returns. A decile spread portfolio that sells stocks with high disagreement and buys stocks with low disagreement earns a value-weighted return of 13% per year. Further analyses suggest MFD-alpha is mispricing induced by short-sale costs and limits-to-arbitrage. (*JEL* G10, G11, G12, G14)

Received: May 30, 2024
Editor: Clemens Sialm
Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

We thank Clemens Sialm and two anonymous referees for their insightful and constructive comments. We also benefited from discussions with Tarun Chordia, Ian Dew-Becker, Victor DeMiguel, Stefano Giglio, Paul Goldsmith-Pinkham, Christian Goulding, Benjamin Holcblat, Narasimhan Jegadeesh, Theis Ingerslev Jensen, Wei Jiang, Paul Karehnke, Ali Kakhbod, Axel Kind, Alejandro Lopez-Lira, David McLean, Roni Michaely, Andreas Neuhierl, Lasse Pedersen, Paul Schneider, Kelly Shue, Taisiya Sikorskaya, Petra Sinagl, Xiaofei Zhao, and seminar participants at the SFS Cavalcade North America Annual Meeting 2025, WFA Annual Meeting 2025, 1st Alpine Finance Conference, 4th Annual Bristol Financial Markets Conference, 21st Annual Bernstein Quantitative Finance Conference, the Inaugural FutureFinTech Federated Conference at the University of Luxembourg, University of Washington 6th Summer Finance Conference, Emory University, ESCP Paris, Georgetown University, Heinrich-Heine-University Duesseldorf, Norwich Business School, Yale School of Management, Technical University of Munich, and the University of Konstanz. We are grateful to Jianan Liu for providing the Stambaugh, Yu, and Yuan (2015) mispricing scores. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Bryan T. Kelly, bryan.kelly@yale.edu.

The Review of Financial Studies 00 (2026) 1–51

https://doi.org/10.1093/rfs/hhag042 Advance Access publication April 21, 2026

Belief disagreement is a primary motivation for trade; thus, understanding disagreement is critical to understanding the behavior of financial markets. A large theoretical literature explores how belief heterogeneity can persist even when agents have access to the same information. Harris and Raviv (1993) argue that differences of opinion among rational investors, even with common information, can drive trading activity and price dynamics.

Other models show that when investors interpret shared signals differently or face frictions, such as short-sale constraints, a difference of opinion can persist and meaningfully influence asset prices and trading behavior. A prominent example is Miller (1977), who predicts that stock prices are biased upward when there is a divergence of opinions among investors about stock value and pessimistic investors face short-sale constraints.[1] Disagreement about public signals has also been linked to trading volume dynamics (Kandel and Pearson 1995; Banerjee and Kremer 2010) and return predictability (Allen, Morris, and Shin 2006).

Despite the rich theoretical literature, empirical work on disagreement is more limited because of the difficulty in measuring investor beliefs. Beliefs are inherently unobservable, and disagreement must often be indirectly inferred through proxies. A seminal study by Diether, Malloy, and Scherbina (2002) proxies for belief heterogeneity using data on equity research analyst earnings forecasts. They show that higher analyst forecast dispersion (AFD) predicts lower future return in the cross section of individual stocks. Johnson (2004) questions the interpretation of Diether, Malloy, and Scherbina (2002) and argues that AFD proxies for firm-specific risk.[2] This paper addresses the challenge of measuring disagreement by empirically simulating it: investors share the same information set but are each assigned a different machine learning model to process it, allowing us to generate belief dispersion in a transparent and theoretically grounded way.

In this paper, we propose a new measure of investor disagreement. We then investigate whether our measure predicts cross-sectional variation in future returns of individual stocks in general and also in light of heterogeneous short-sale frictions. We contribute to the literature in three ways.

Our first contribution is introducing a new measure of belief disagreement at the asset level. Because the distribution of investor beliefs is not directly observable, we propose a statistical surrogate. Each investor is

[1] Other important theoretical contributions include Harrison and Kreps (1978), Hong and Stein (1999), Chen, Hong, and Stein (2002), Hong and Stein (2003), Scheinkman and Xiong (2003), and Atmaz and Basak (2018).

[2] Other empirical papers studying the effect of belief disagreement include Anderson, Ghysels, and Juergens (2005), Barber and Odean (2008), Yu (2011), and Jiang and Sun (2014).

a prediction model from which beliefs about future returns are formed. These hypothetical investors have access to a common set of predictive information, but different investors use available information in different ways. We simulate the distribution of beliefs by endowing each investor with a machine learning model but introduce random variation in model specification across investors. By randomizing the set of model specifications, we capture the idea that investors have a distribution of prior beliefs, information frictions, and biases. Yet all investors in our model are sophisticated (à la Stein 2009), though imperfect, optimizers. They are sophisticated in the sense that each investor's model is a random forest model that uses large predictor sets in flexible and nonlinear ways. But they are imperfect in that no investor has a correctly specified model; instead they have a variety of models that are heterogeneous approximations of the true data generating process. They are optimizers in that, given their model endowment, each investor uses the available data to estimate model parameters and form predictions.

Our approach of measuring disagreement via the dispersion of forecasts from a set of statistical models is conceptually related to the recent work of Avramov et al. (2023), who also quantify model disagreement to study investor heterogeneity. However, our work makes a distinct and complementary contribution. While Avramov et al. (2023) focus on the time-series properties of their disagreement measure and its role in explaining perceived volatility, our primary motivation is to utilize our disagreement measure as a powerful tool to simulate differences in beliefs of investors and as cross-sectional signal for predicting future stock returns.

A key question is whether the distribution from which we simulate model specifications (and hence investors' belief formation processes) is plausible. A significant portion of our analysis is dedicated to this question. We argue that (a) the calibration of our simulation distribution is reasonable and (b) our results are robust to a range of distributions for simulating investors' models.

Given our construction of investor beliefs, we then measure stock-level disagreement as dispersion in investors' future return (or earnings) forecasts, which we refer to as machine forecast disagreement (MFD). MFD has several attractive attributes. By sidestepping the difficult problem of directly and reliably surveying investor beliefs, the data coverage of MFD is much better than prior literature, which is essentially constrained by the availability of analyst forecasts from I/B/E/S. In contrast, MFD is available for all stocks at all times. Also, MFD is arguably a more objective measure of disagreement than AFD. While analysts are undoubtedly important information intermediaries in financial markets, evidence points to biases in their recommendations driven, for example, by incentives to secure underwriting and other investment banking business (see, e.g., Dugar and Nathan 1995; Michaely and Womack 1999; Chan, Karceski, and

Lakonishok 2007). While we argue that machine learning models suffer less from behavioral biases or conflicts of interest, one may counter that our distribution of model specifications is biased in other ways. An attractive feature of MFD is that it constitutes a complete methodology for modeling and measuring disagreement. Shortcomings or biases in our specific implementation can be reformulated by other researchers to incorporate richer and more realistic belief simulations, and our results can in turn be reanalyzed in light of such model improvements.

Our second main contribution is documenting the strong predictive power of MFD for the cross-sectional pricing of individual stocks. We find that stocks with higher MFD earn significantly lower future returns than otherwise similar stocks. In particular, a value-weighted (equal-weighted) portfolio of stocks in the highest MFD decile underperforms a portfolio of stocks in the lowest MFD decile by 1.09% (1.33%) per month with a Newey-West t-statistic of 4.32 (5.64). We also present extensive evidence that validates MFD as an effective measure of belief disagreement. While MFD has on average a 23% cross-sectional correlation with analyst disagreement, AFD is less correlated to alternative measures of investor disagreement compared to MFD. Moreover, AFD is a notably weaker predictor of stock returns. The analogous value-weighted (equal-weighted) return of an AFD-based portfolio is 0.42% (0.79%) per month with a t-statistic of 1.58 (3.75). In Fama-MacBeth regressions, MFD is among the most statistically significant predictors of returns after controlling for other commonly studied characteristics, including value, investment, profitability, momentum, reversal, illiquidity, and volatility. We also show that the cross-sectional return prediction power of MFD extends to international equity markets (excluding the United States) with the magnitudes and significance of international prediction effects closely in line with those for our main U.S. sample.

Our methodology remarkedly distinguishes between different sources of investor disagreement. We construct our primary measure, MFD, based on forecasts of future returns. We also construct an alternative measure based on forecasts of future earnings. This dual approach naturally maps onto the fundamental asset pricing decomposition, where expected returns reflect both expected discount rates and expected cash flows. Disagreement over future returns can be interpreted as heterogeneity in beliefs about discount rates, while disagreement over future earnings relates to heterogeneity in beliefs about future cash flows. MFD's effect on future stock returns remains strong if it's based on realized next-quarter earnings instead of future excess returns. Moreover, as expected, the cross-sectional correlation between analyst disagreement and the earnings-based measures of MFD is stronger, compared to the return-based measure of MFD.[3]

Our third contribution is investigating the economic underpinnings of MFD alpha. First, we condition our analysis on short-sale constraints. The overpricing of high-MFD stocks is especially pronounced among stocks with high short-sale costs. Stocks in the highest tercile of indicative borrowing fees experience an alpha spread of −2.36% per month versus −0.12% for stocks in the lowest tercile. The difference of −2.24% (t-stat. = −6.06) is strongly supportive of the hypothesis that disagreement results in assets being more overpriced in the presence of more binding short-sale constraints. We find similar supportive evidence based on institutional ownership. The alpha spread on MFD-sorted portfolios of stocks with high retail ownership is −1.10% per month, lower than the alpha spread on MFD-sorted portfolios of stocks largely held by institutional investors of −0.31% per month. The difference between these two alpha spreads, 0.79% per month (t-stat.= 4.53), is highly significant and further supports the Miller (1977) hypothesis.

Next, we find supportive evidence that the MFD premium is associated with high-MFD stocks being mispriced, measured by the stock-level mispricing (MISP) definition of Stambaugh, Yu, and Yuan (2015). We find an MFD alpha spread of −0.68% per month for stocks in the highest MISP tercile (i.e., overvalued stocks), compared to a spread of −0.31% for stocks in the lowest MISP tercile. The difference of these alpha spreads is statistically significant (−0.37% with t-stat. = −2.51), suggesting that high-MFD stocks have a significantly higher mispricing score (or higher degree of overpricing) than low-MFD stocks.

We document additional support for the interpretation that MFD alpha is driven by stock-level mispricing by examining stock price reactions around earnings announcements. Assuming that investors exhibit biased expectations and are overly optimistic about high-MFD stocks, they update their beliefs in the presence of new information leading to a stock price correction (Engelberg, McLean, and Pontiff 2018). Hence, the return prediction during earnings announcements should exceed that of nonearnings periods. In line with this intuition, the return spread for the hedged MFD strategy is 156% (113%) higher during a 1-day (3-day) earnings announcement window than on nonannouncement days. We also find that the MFD alpha is significantly stronger for stocks with more severe limits-to-arbitrage, consistent with limits-to-arbitrage exacerbating asset mispricing.

[3] The average cross-sectional correlations between AFD and the earnings-based measures of MFD are in the range of 26% and 53%, while the average cross-sectional correlation between AFD and the return-based measure of MFD is 23%.

1. An Empirical Model of Disagreement

1.1 Modeling investor beliefs

Gu, Kelly, and Xiu (2020) consider a general conditional risk premium formulation:

$$E_t[r_{i,t+1}]=g(z_{i,t}),$$

where $z_{i,t}\in\mathbb{R}^d$ is data comprising the time t information set about asset i that is available to investors, and $g(\cdot)$ is a general (likely nonlinear) function mapping that information into risk premiums.

To model disagreement, we consider a collection of investors $k=1,...,K$. Each investor k differs in her information set $z_{k,i,t}$ and how she forms expectations based on $z_{k,i,t}$. In particular, an investor k forms beliefs according to

$$E_{k,t}[r_{i,t+1}]=g_k(z_{k,i,t}). \tag{1}$$

That is, investor beliefs can disagree.

The idea that investors disagree is uncontroversial. As outlined by Barberis (2018), disagreement lies at the heart of many behavioral models of financial markets and is critical for generating the large trading volumes observed in many markets. Although the precise sources of disagreement are not well-understood, "if two people are to disagree, one of three things must be true: (1) they have different prior beliefs; (2) they observe different information; or (3) one or both of them is not fully rational" (Barberis 2018).

We propose a belief-generating model from which we build an empirical measure of investor disagreement. In particular, we simulate differences in beliefs across investors by endowing them with different models for forecasting returns from investor-specific inputs $z_{k,i,t}$. We assume investor k forecasts returns according to

$$g_k(z_{i,t})=RF_k(z_{k,i,t}), \tag{2}$$

where $RF_k(\cdot)$ denotes random forest regression. The investor-specific beliefs in Equation (2) have two main components that drive disagreement. In the first component, each investor is endowed with an incomplete information set, $z_{k,i,t}\in\mathbb{R}^{d_k}, 1\leq d_k\leq d$, of the common input data, $z_{i,t}\in\mathbb{R}^d$. That is, each investor has access only to a subset l_k of the entire d-dimensional information set $z_{i,t}$. Let $z^l_{i,t}\in\mathbb{R}$ denote the $l-$th dimension of $z_{i,t}$. Then the investor-specific information set is given as,

$$z_{k,i,t}=[z^{l_{k,1}}_{i,t},...,z^{l_{k,d_k}}_{i,t}], \tag{3}$$

where $l_k=\{l_{k,i}\in\{1,...,d\}|\forall i\neq j:l_{k,i}\neq l_{k,j}\}$ and $|l_k|=d_k$. Each investor k accesses the common data in her own idiosyncratic way, transforming it into a feature set that is unique to k (though, naturally, correlated with other investors' views as well). These features summarize investor k's perception of the world around her, and can be interpreted as capturing differences in investors' access to information, information processing ability, or perceptive biases. Note that the dimension of the investor-specific information might vary across investors k. To illustrate this, consider a simple case where the complete information set $z_{i,t}$ includes past returns and fundamental and accounting-based ratios. Investor A might focus primarily on technical indicators, accessing only past returns (e.g., $r_{i,t-1},...,r_{i,t-12}$) and volatility. Investor B might be fundamentally oriented, focusing on the price-to-book ratio and dividend yield. In case of a strong decline of a stock, investor A expects future negative returns, whereas investor B believes in positive future returns as she views the same stock as fundamentally strong having declined below her estimate of fair value. Hence, heterogeneity in information sets naturally leads to different forecasts and beliefs even when investors are rational and use optimal forecasting techniques.

In the second stage, investors estimate $g_k(\cdot)$ using a random forest regression. Our motivation for this component is twofold. First, the regression represents optimizing behavior on the part of investors as they learn how to best use their individual feature sets in a flexible specification. Random forest regression is known to be effective for capturing nonlinearities and interaction effects in financial forecasting problems with high-dimensional predictor sets (see Gu, Kelly, and Xiu 2020; Van Binsbergen, Han, and Lopez-Lira 2023, for applications to return and earnings prediction, respectively). Second, random forest introduces a further layer of heterogeneity in investor beliefs by randomizing regression specifications across investors (through the use of bootstrapping and dropout), which can be interpreted as heterogeneous model priors across investors.

In summary, our specification of $g_k(z_{i,t})$ is a reduced-form representation that accommodates the three potential sources of disagreement outlined by Barberis (2018): (1) different prior beliefs are captured through the random nature of forest construction, where each investor's model represents a different prior about the relationship between information and future returns; (2) different information sets are explicitly modeled by providing each investor with a unique subset of the complete information set; and (3) varying rationality is implicitly captured through differences in model complexity and feature emphasis in random forests, where some investors' models may be more sensitive to relevant predictive signals than others.

Once investors are endowed with a model and estimate the model subject to their respective data sets, they construct return forecasts for each stock in each month, $E_{k,t}[r_{t+1}]$. We measure disagreement, MFD, for stock i as the standard deviation of $E_{k,t}[r_{i,t+1}]$ across investors:

$$MFD_{i,t} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(E_{k,t}[r_{i,t+1}] - \overline{E_t[r_{i,t+1}]}\right)^2}, \tag{4}$$

where $\overline{E_t[r_{i,t+1}]} = \frac{1}{K}\sum_{k=1}^{K} E_{k,t}[r_{i,t+1}]$ is the average forecast across all investors.[4] MFD captures the dispersion in beliefs that arises from the heterogeneity in both information sets and model specifications described above. Higher values of MFD indicate greater disagreement among investors about a stock's future returns.

1.2 Motivation for modeling choices

After having described our modeling choice for measuring disagreement at the asset level, we outline the motivation of its implicit and explicit assumptions.

We focus on modeling beliefs of sophisticated investors. Financial markets offer high rewards for success with relatively few barriers to entry. Given this great deal at stake, market participants like investors and money managers are frequently assumed to be sophisticated individuals (Kandel and Pearson 1995). Related, Stein (2009) documents that sophisticated investors are increasingly dominating stock trading.

Sophistication of investors is manifested in three dimensions in our methodology: (1) employing a large set of characteristics, (2) using sophisticated statistical models allowing for nonlinearities and interaction effects between predictor variables, and (3) optimizing behavior. Li and Rossi (2020) show that the majority of mutual funds are exposed to 40 to 50 stock characteristics which constitute approximately 50% of the characteristics the authors use in their study. This implies that the information set in Equation (3) is highly dimensional for each investor. Moreover, findings in Li and Rossi (2020) support the choice of using random forests in the expectation formulation function in Equation (2). Li and Rossi (2020) document that the performance of funds is nonlinearly related to characteristics of the stocks they hold, there exist significant interactions between various stock characteristics at the mutual fund level and fund performance, and that tree-based ensemble models predict fund performance most accurately. O'Doherty, Savin, and Tiwari (2017) complement the findings of Li and Rossi (2020) by showing that strategies of hedge funds, which are usually considered the most sophisticated investors, are best modeled with a combination approach of simple linear

[4] We define MFD as the equal-weighted standard deviation of forecasts across models. This differs slightly from Avramov et al. (2023), who use a probability-weighted formulation in their analysis.

models, similar to the ensemble approach in random forests, employed in Equation (2).

Moreover, random forest regression in Equation (2) implies the use of statistical models and optimizing behavior in the formation of beliefs. Andries et al. (2025) provide experimental evidence for our assumption. The authors show that investors make rational forecasts when they receive predictive signals, while investors use extrapolative expectations when no useful information is provided. As asset characteristics contain predictive power for future returns (Gu, Kelly, and Xiu 2020) and form the basis for beliefs in our setting, the findings in Andries et al. (2025) directly motivate our choice of modeling sophisticated investors as being optimizers.

Dahlquist and Ibert (2024) and Couts, Gonçalves, and Loudis (2024) complement the findings of Andries et al. (2025) with respect to subjective beliefs of institutional investors. Both papers utilize capital market assumptions of major asset managers and institutional investor consultants to extract subjective expected returns at the asset-class level. Subjective beliefs of these market participants largely match objective (data driven), statistical beliefs. These findings highlight that the use of a statistical surrogate for belief formation is sensible.

Moreover, Dahlquist and Ibert (2024) and Couts, Gonçalves, and Loudis (2024) provide evidence in favor of two additional assumptions of our approach. First, both papers document considerable heterogeneity in subjective expectations of asset managers. Dahlquist and Ibert (2024), for example, show that the cross-sectional standard deviation of asset managers' subjective expectations is 73% larger than the time-series standard deviation and 78% of the variation are explained by manager fixed effects.[5] Couts, Gonçalves, and Loudis (2024) document similar persistence. Persistence of heterogeneity aligns well with our assumption that no single investor has the "right model", but investors are "imperfect" optimizers given their information set and prediction model. Finally, Couts, Gonçalves, and Loudis (2024) document that the predictability of subjective expected returns for future realized returns is, to a large extent, driven by subjective risk premiums instead of subjective alphas, which stresses the necessity of strong subjective risk premiums rather than subjective alphas when modeling subjective return expectations. Equation (1) directly incorporates this finding.

1.3 Benefits

MFD comes with various advantages over existing survey-based measures of disagreement, such as AFD. First, MFD can be constructed for many more U.S. stocks across a longer time horizon. In our setting, we can

[5] The persistent heterogeneity in beliefs has also been documented for retail investors (Giglio et al. 2021; Laudenbach et al. 2024).

construct MFD for about 67% more stocks on average compared to AFD for the time period for which AFD is available. Second, a common concern regarding survey-based measures is the relatively short time horizon of coverage. MFD overcomes this in that it only relies on stock-level characteristics to build the measure and can be calculated before analyst coverage became available; I/B/E/S has been mainly used after 1983, whereas data on firm fundamentals date back to at least 1950 from easy-to-access databases. Third, and related to the previous point, survey-based measures are typically not found for international stocks; MFD can cover international stocks, as long as characteristics data are available for these stocks. This might be of particular relevance for countries with little to no analyst coverage. In general, our measure of disagreement can be applied to any asset type, for example, commodities, currencies, or credit instruments, if asset characteristics are available. Moreover, as MFD does not rely on the update cycle of analyst recommendations, it can be updated at the discretion of the researcher.

An additional benefit of MFD is the flexibility of the belief simulation mechanism. While we argue that MFD could be less prone to the behavioral biases or conflicts of interest found in survey responses (see, e.g., Dugar and Nathan 1995; Michaely and Womack 1999; Chan, Karceski, and Lakonishok 2007), our disagreement mechanism might lack certain features in the view of other researchers. Belief simulation is easily modifiable in our methodology. For example, extrapolative (Da, Huang, and Jin 2021; Nagel and Xu 2023) despite countercyclical beliefs (Dahlquist and Ibert 2024; Couts, Gonçalves, and Loudis 2024) can be incorporated by putting more weight on recent observations.

The easy-to-change, modular bottom-up approach of modeling investor beliefs yields further benefits, especially over previously established disagreement proxies like idiosyncratic volatility (Boehme, Danielsen, and Sorescu 2006; Berkman et al. 2009) or trading volume (Boehme, Danielsen, and Sorescu 2006; Garfinkel 2009; Banerjee 2011). These measures are outcome variables that might be caused by disagreement, but could also proxy for other phenomena beyond disagreement. MFD measures disagreement at its source. Furthermore, it can be decomposed into its constituent sources in ways that measures like idiosyncratic volatility cannot. For example, it allows us to better understand whether disagreement stems from access to differing information sets or from different ways to interpret information.[6] Or it could be used to identify which type of information yields the highest disagreement. Additionally, our approach to modeling disagreement can provide valuable empirical insights into the maximum attainable levels of investor belief dispersion, offering useful guidance

[6] We perform this decomposition in Section 5.6.

for theoretical asset pricing models that incorporate parameters for belief heterogeneity.

Finally, some of the previously proposed measures of disagreement, like idiosyncratic volatility, are backward-looking. In the case of MFD, however, even though investors optimize their models on past data, their predictions are inherently forward-looking and capture expectations about future performance, much like survey-based beliefs.

2. Data and Variables

We use the data set from Jensen, Kelly, and Pedersen (2023), a publicly available data set of stock returns and characteristics.[7] The underlying return data are sourced from the Center for Research in Security Prices (CRSP) and accounting data from Compustat.[8] We restrict our sample to common stocks trading at the NYSE, AMEX, and NASDAQ. We exclude financial and utilities firms. To reduce the effect of small and illiquid stocks, we also exclude low-priced stocks trading below $5 per share.

To predict returns, we use the 153 stock characteristics as the complete information set z. We cross-sectionally rank all stock characteristics period-by-period and map these ranks onto the $[-1,1]$ interval following Kelly, Pruitt, and Su (2019), Gu, Kelly, and Xiu (2020), and Freyberger, Neuhierl, and Weber (2020).

Our sample covers the period from July 1966 to December 2022. Our approach utilizes a 10-year rolling window to estimate the random forest regressor. We calculate the month-t MFD using characteristics from the previous month $(t-1)$. Subsequently, we conduct out-of-sample cross-sectional asset pricing tests for the period August 1976 to December 2022.

2.1 MFD construction

The specific procedure for constructing our measure of disagreement, MFD, is as follows. We set the number of investors, K, to 100, and the dimension of the incomplete information set to 76, that is, $d_k = 76$ for all $k = 1,...,K$. This mimics the findings in Li and Rossi (2020) showing that mutual funds are exposed to 40–50 characteristics of stocks they hold, constituting 50% of the stock characteristics the authors consider. The random forest regression model has several hyper-parameters. Our baseline choices, which are standard and similar to those in Gu, Kelly, and Xiu (2020), are described in Table B1 in the appendix. We confirm that our

[7] The data, replication code, and documentation can be found at https://github.com/bkelly-lab/ReplicationCrisis/tree/master/GlobalFactors.

[8] In principle, our methodology allows for any type of stock characteristics, such as option-derived characteristics (Neuhierl et al. 2025). We focus on characteristics derived from CRSP and Compustat data as these sources are academic standard and allow for a long sample period.

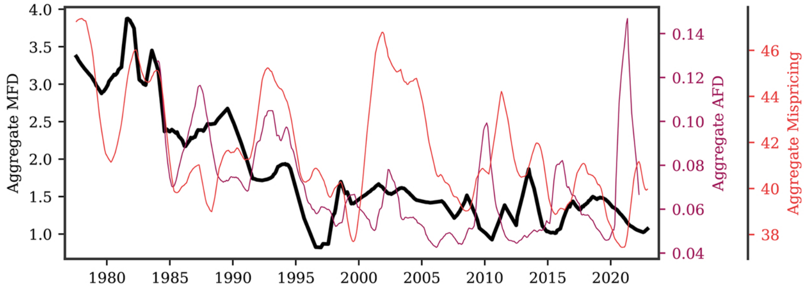

Figure 1
Aggregate MFD over time
The figure shows the 12-month rolling average of value-weighted aggregate MFD (annualized, in percent) in black on the right y-axis. On the left y-axis, the figure depicts the 12-month rolling average of value-weighted aggregate analyst forecast dispersion (AFD; Diether, Malloy, and Scherbina 2002) and stock mispricing index of Stambaugh and Yuan (2017). The sample period is from August 1976 to December 2022.

results are robust to a variety of alternative hyperparameter specifications and using gradient boosted regression trees (Friedman 2001) in Equation (2) in Section 5.

2.2 Descriptive statistics

Table 1 presents summary statistics for the main cross-sectional variables. In our cross-sectional regression analysis, we control for a comprehensive set of firm characteristics known to predict future returns. Detailed definitions and the original sources for all control variables are provided in Table A1 in the appendix. Concerning our key variable of interest, MFD, the time-series average of the annualized cross-sectional mean is 1.93% with an average cross-sectional standard deviation of 0.53%. The average annualized cross-sectional 10$^{\text{th}}$ percentile of MFD is 1.31%, while the 90$^{\text{th}}$ percentile is 2.5%, indicating a positively skewed distribution of MFD.

Figure 1 displays the annual time-series plot of the aggregate MFD. Aggregate value-weighted MFD is higher (lower) when analysts disagree more (less) and when there is more (less) overvaluation in the equity market. The correlation of aggregate MFD to aggregate AFD is 36%, whereas it is 48% to an aggregate mispricing score based on the stock-level mispricing factor of Stambaugh, Yu, and Yuan (2015).

Table 2 includes the cross-sectional Spearman's rank correlation coefficient of MFD to the aforementioned control variables. The first column and the first row report a negative relation between MFD and 1-month-ahead returns in excess of the risk-free rate. It further shows that smaller and less liquid stocks with higher analyst dispersion and higher idiosyncratic volatility also have higher MFD. This positive (negative) correlation of MFD with idiosyncratic volatility (size) suggests that the machine forecast

Table 1
Descriptive statistics

	Mean	Sd	10^{th}	Q1	Q2	Q3	90^{th}
RET_{t+1}	0.77	12.94	−12.77	−6.08	0.11	6.71	14.66
MFD	1.93	0.53	1.31	1.52	1.91	2.27	2.65
SUE	−0.08	1.77	−1.89	−0.76	0.01	0.77	1.77
AG	0.29	0.74	−0.07	0.01	0.10	0.26	0.70
MOM	0.23	0.53	−0.27	−0.09	0.12	0.40	0.82
ILLIQ	0.91	2.60	0.00	0.02	0.10	0.55	2.19
OP	0.22	0.39	−0.07	0.11	0.23	0.34	0.50
IVOL	0.03	0.01	0.01	0.02	0.03	0.03	0.04
BETA	1.20	0.62	0.48	0.77	1.12	1.54	2.04
SIZE ($\times 10^{-9}$)	3.74	17.09	0.07	0.17	0.51	1.73	6.12
BM	0.62	0.46	0.16	0.29	0.51	0.83	1.22
MAX	0.04	0.02	0.02	0.02	0.03	0.05	0.06
TURN ($\times 10^{3}$)	6.25	9.82	1.24	2.51	4.44	7.41	12.00
STR	0.02	0.12	−0.11	−0.05	0.01	0.08	0.16
AFD	0.16	0.42	0.01	0.02	0.05	0.11	0.31

The table reports the summary statistics for the cross-sectional variables. The sample consists of all common stocks that are listed on NYSE, Amex, and Nasdaq. Financial firms (with one-digit SIC = 6), utility firms (with two-digit SIC = 49), and stocks trading below \$5/share are excluded from the analysis. RET_{t+1} is the 1-month-ahead return in excess of the risk-free rate of individual stocks. MFD is the machine forecast disagreement variable. SIZE is the firm's market capitalization at the end of month $t-1$ (Fama and French 2008). BM is the ratio of the firm's book value of equity divided by its market capitalization, following Fama and French (2008). Asset growth (AG) is a percentage of total asset growth between 2 consecutive fiscal years, following Cooper, Gulen, and Schill (2008). Operating profits (OP) is the ratio of operating profits to book equity, following Fama and French (2015). Short-term reversal (STR) is the stock's 1-month-lagged return, following Jegadeesh (1990). MOM is the stock's cumulative return from the start of month $t-12$ to the end of month $t-1$, skipping STR, following Jegadeesh and Titman (1993). ILLIQ is the Amihud (2002) illiquidity measure computed using daily data over the last 126 trading days. TURN is the share turnover computed over the last 126 trading days, following Datar, Naik, and Radcliffe (1998). SUE is the standardized unexpected earnings, following Foster, Olsen, and Shevlin (1984). IVOL is the standard deviation of daily residuals estimated from the daily regression of excess stock returns on the excess market return over the previous year, following Ali, Hwang, and Trombley (2003). MAX is the average of the five highest daily returns of each stock in month $t-1$, following Bali, Cakici, and Whitelaw (2011). All variables are winsorized at the 1% level for both tails to mitigate the effect of outliers. The mean, standard deviation (Sd), 10th percentile (10th), first up to third quartile, and the 90th percentile (90th) are shown. The sample is from August 1976 to December 2022.

disagreement is also a reasonable proxy for information uncertainty (see, e.g., Johnson 2004; Zhang 2006).

3. Empirical Results

In this section, we conduct parametric and nonparametric tests to assess the predictive power of machine forecast disagreement (MFD) over future stock returns. First, we present results of the univariate portfolio-level analysis. Second, we report the average stock characteristics of the MFD-sorted decile portfolios. Third, we conduct bivariate portfolio-level analyses to assess the predictive power of MFD after controlling for well-known stock characteristics and risk factors. Finally, we present firm-level Fama-MacBeth cross-sectional regression results.

Table 2
Cross-sectional correlations to MFD

	Mean	Sd	10^{th}	Q1	Q2	Q3	90^{th}
RET_{t+1}	−0.05	0.09	−0.17	−0.10	−0.04	0.01	0.07
SUE	−0.02	0.08	−0.13	−0.08	−0.02	0.02	0.08
AG	0.13	0.15	−0.08	0.03	0.13	0.24	0.32
MOM	−0.07	0.17	−0.26	−0.18	−0.10	0.05	0.19
ILLIQ	0.14	0.22	−0.22	0.04	0.20	0.30	0.38
OP	−0.29	0.16	−0.48	−0.43	−0.34	−0.13	−0.10
IVOL	0.40	0.30	−0.20	0.33	0.52	0.60	0.65
BETA	0.19	0.10	0.04	0.12	0.20	0.25	0.32
SIZE	−0.15	0.27	−0.39	−0.34	−0.25	−0.07	0.41
BM	−0.05	0.16	−0.27	−0.16	−0.05	0.08	0.14
MAX	0.35	0.22	−0.06	0.31	0.43	0.50	0.55
TURN	0.16	0.12	0.01	0.09	0.16	0.23	0.31
STR	0.01	0.11	−0.12	−0.05	0.02	0.08	0.14
AFD	0.21	0.17	−0.00	0.09	0.23	0.35	0.41

The table reports summary statistics on the cross-sectional correlations of various stock characteristics with MFD. Correlation is measured using Spearman's ρ. The stock characteristics are defined in Table 1. The mean, standard deviation (SD), 10th percentile (10th), first, second, and third quartiles, and the 90th percentile (90th) are shown. The sample is from August 1976 to December 2022.

3.1 Univariate portfolio-level analysis

To construct the long-short portfolio for each month from August 1976 to December 2022, individual stocks are sorted by MFD into decile portfolios. We then compute the 1-month-ahead value-weighted and equal-weighted average excess return of each decile portfolio. To examine the cross-sectional relation between MFD and future stock returns, we form a long-short portfolio that takes a long position in the lowest decile of MFD and a short position in the highest decile of MFD.

In Table 3, we report the average monthly excess returns (in excess of the 1-month Treasury-bill rate) of each decile portfolio, and the long-short portfolio. We also analyze abnormal returns (alphas) using different factor models. These include the capital asset pricing model (CAPM) with the market factor (MKT), the six-factor model (FF6) by Fama and French (2018) including MKT, size (SMB), value (HML), investment (CMA), profitability (RMW), and momentum (MOM) factors. Furthermore, we use the q4-factor model (HXZ) by Hou, Xue, and Zhang (2015) with MKT, size (SMB_Q), investment (I/A), and profitability (ROE) factors. We also consider the mispricing factor model (SY) of Stambaugh and Yuan (2017) with MKT, SMB, management (MGMT), and performance (PERF) factors, along with the behavioral factor model (DHS) of Daniel, Hirshleifer, and Sun (2020) using MKT, post-earnings-announcement drift (PEAD), and financing (FIN) factors.

In general, the excess returns and the alphas of the MFD-sorted portfolios decrease from decile 1 to decile 10. The long-short portfolio that short-sells stocks in the highest 10th percentile of MFD (decile 10) and buys stocks in the lowest 10th percentile of MFD (decile 1) earns a value-weighted (equal-weighted) average return of 1.09% (1.33%) per

Table 3
Univariate portfolio sorts on MFD

A. Equal-weighted portfolios

	Excess return	t-stat	CAPM	t-stat	FF6	t-stat	HXZ	t-stat	SY	t-stat	DHS	t-stat
Low	1.14***	(5.10)	0.50***	(3.69)	0.23***	(2.72)	0.27***	(2.74)	0.28***	(2.69)	0.50***	(3.74)
2	1.08***	(4.87)	0.43***	(3.35)	0.21***	(3.33)	0.25***	(3.16)	0.22***	(2.77)	0.45***	(3.62)
3	1.03***	(4.50)	0.34***	(2.69)	0.16**	(2.45)	0.22***	(2.86)	0.17**	(2.31)	0.44***	(3.42)
4	1.01***	(4.36)	0.30**	(2.47)	0.15***	(2.91)	0.22***	(3.23)	0.17**	(2.46)	0.42***	(3.42)
5	0.97***	(4.03)	0.23*	(1.96)	0.14***	(2.69)	0.23***	(3.97)	0.15**	(2.57)	0.42***	(3.63)
6	0.84***	(3.34)	0.08	(0.68)	0.05	(1.14)	0.13**	(2.35)	0.03	(0.50)	0.31***	(2.70)
7	0.79***	(2.93)	−0.01	(−0.11)	0.03	(0.54)	0.16***	(2.86)	0.05	(0.71)	0.32**	(2.50)
8	0.60**	(2.08)	−0.23	(−1.56)	−0.04	(−0.67)	0.10	(1.33)	0.00	(0.04)	0.28**	(1.99)
9	0.45	(1.45)	−0.43***	(−2.65)	−0.14**	(−2.02)	0.05	(0.60)	−0.11	(−1.10)	0.22	(1.42)
High	−0.19	(−0.55)	−1.13***	(−5.75)	−0.67***	(−8.01)	−0.48***	(−3.89)	−0.72***	(−5.13)	−0.29	(−1.64)
H-L	−1.33***	(−5.64)	−1.63***	(−7.25)	−0.90***	(−6.71)	−0.76***	(−4.25)	−1.00***	(−4.80)	−0.78***	(−4.66)

B. Value-weighted portfolios

	Excess return	t-stat	CAPM	t-stat	FF6	t-stat	HXZ	t-stat	SY	t-stat	DHS	t-stat
Low	0.91***	(4.73)	0.33***	(3.15)	0.04	(0.46)	0.04	(0.39)	0.09	(0.84)	0.07	(0.68)
2	0.90***	(4.40)	0.27***	(3.22)	0.13	(1.60)	0.11	(1.28)	0.13	(1.46)	0.12	(1.33)
3	0.92***	(4.53)	0.25**	(2.41)	0.15	(1.46)	0.17*	(1.66)	0.14	(1.27)	0.20**	(1.97)
4	0.86***	(4.14)	0.16**	(2.11)	0.22***	(2.90)	0.21**	(2.53)	0.25***	(2.69)	0.24***	(3.09)
5	0.75***	(3.53)	0.04	(0.54)	−0.02	(−0.24)	−0.03	(−0.40)	0.02	(0.30)	0.06	(0.75)
6	0.72***	(3.17)	0.01	(0.07)	−0.02	(−0.19)	0.01	(0.09)	−0.08	(−0.83)	0.03	(0.35)
7	0.75***	(2.96)	−0.04	(−0.43)	0.00	(0.03)	0.03	(0.28)	−0.02	(−0.17)	0.09	(0.98)
8	0.44*	(1.86)	−0.32**	(−2.56)	−0.09	(−0.82)	−0.06	(−0.49)	−0.02	(−0.16)	0.04	(0.40)
9	0.46	(1.48)	−0.40***	(−2.67)	−0.13	(−0.92)	0.02	(0.13)	−0.07	(−0.44)	0.04	(0.29)
High	−0.18	(−0.55)	−1.09***	(−5.67)	−0.55***	(−3.95)	−0.40**	(−2.33)	−0.43***	(−2.76)	−0.45***	(−3.13)
H-L	−1.09***	(−4.32)	−1.42***	(−5.82)	−0.60***	(−3.35)	−0.44**	(−2.17)	−0.51**	(−2.51)	−0.52***	(−2.93)

The table reports the average monthly excess returns and alphas on univariate portfolios of stocks sorted by MFD. Each month t, stocks are sorted into decile portfolios by MFD constructed using data up to month $t-1$. Panel A reports equal-weighted portfolio sorts, whereas panel B reports value-weighted portfolio sorts. Excess return is the return in excess of the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models: the CAPM, Fama and French (2018) six-factor model (FF6), Stambaugh and Yuan (2017) mispricing factor model (SY), Hou, Xue, and Zhang (2015) q-factor model (HXZ), and the Daniel, Hirshleifer, and Sun (2020) behavioral factor model (DHS). t-stat refers to Newey and West (1987) adjusted t-statistics. The sample period is from August 1976 to December 2022 (December 2016 in case of SY). $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

month with a t-statistic of 4.32 (5.64), translating into an annualized return of 13.08% (15.96%).[9] Controlling for the robust risk and mispricing factors does not change the magnitude and statistical significance of the return spreads on the MFD-sorted portfolios for most of the factor models.[10] Notably, the negative association between MFD and future returns is more concentrated in the short leg of the arbitrage portfolio. The alphas are statistically significantly negative and large in absolute terms among the stocks in decile 10 across all factor models for the value-weighted portfolios. On the contrary, the alphas for all factor models except the CAPM are not statistically significantly different from zero for stocks in decile 1 of the value-weighted portfolios.[11] This suggests that high-MFD firms are overvalued relative to firms with lower MFD. Hence, the return predictability is potentially driven by mispricing rather than compensation for risk.

3.2 Average portfolio characteristics

We investigate whether other firm characteristics can explain the negative relation between MFD and future stock returns. We sort stocks by MFD into decile portfolios each month and calculate the time-series averages of the cross-sectional medians of various firm-specific characteristics for each decile. Table 4 presents the average stock characteristics of each MFD-sorted decile portfolio and the long-short portfolio. The characteristics include the machine forecast disagreement (MFD), log market capitalization (SIZE), log book-to-market ratio (BM), asset growth (AG), operating profitability (OP), medium-term stock momentum (MOM), short-term reversal (STR), illiquidity (ILLIQ), turnover (TURN), standardized unexpected earnings (SUE), idiosyncratic volatility (IVOL), market beta (BETA), and lottery demand (MAX).

Earlier studies found that small, illiquid, lottery-like stocks with high idiosyncratic volatility exhibit high information uncertainty (e.g., Zhang 2006; Kumar 2009; Bali, Cakici, and Whitelaw 2011). Consistent with the literature, Table 4 shows that the stocks with higher MFD are indeed

[9] The t-statistics reported in our tables are Newey and West (1987) adjusted with six lags to control for heteroscedasticity and autocorrelation.

[10] As discussed in Section 4, we confirm the robustness of these findings using the nonlinear stochastic discount factors of Chen, Pelger, and Zhu (2024) and Cong et al. (2025).

[11] The significantly positive alpha for the low-MFD portfolio, particularly in the equal-weighted sorts, suggests factors like investor neglect can contribute to the systematic underpricing of these stocks. To test this informational underreaction hypothesis, we conduct an analysis of stock returns around earnings announcements that is analogous to the test for high-MFD stocks in Section 7.1. We find that the positive returns of low-MFD stocks are significantly amplified during the announcement window, consistent with earnings releases catalyzing the correction of misinformation. Moreover, these stocks also exhibit a persistent positive drift on nonannouncement days, suggesting a gradual, ongoing correction of underpricing. These combined findings support a dual-channel underreaction explanation for the performance of the long leg.

Table 4
Average stock characteristics of MFD-sorted portfolios

	Low	2	3	4	5	6	7	8	9	High	H-L	*t*-stat
MFD	1.21	1.39	1.52	1.64	1.84	1.97	2.10	2.27	2.50	2.88	1.67^{***}	(19.42)
SUE	0.03	0.01	−0.00	−0.03	−0.02	−0.01	−0.02	−0.01	−0.01	−0.04	-0.07^{**}	(−2.22)
AG	0.06	0.07	0.08	0.09	0.09	0.11	0.12	0.13	0.16	0.23	0.17^{***}	(7.28)
MOM	0.16	0.13	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.09	-0.07^{**}	(−2.12)
ILLIQ	0.07	0.09	0.07	0.07	0.08	0.09	0.10	0.10	0.11	0.17	0.10^{***}	(4.56)
OP	0.29	0.27	0.26	0.25	0.24	0.23	0.21	0.19	0.15	0.02	-0.27^{***}	(−11.52)
IVOL	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.02^{***}	(21.66)
BETA	0.92	0.98	1.03	1.08	1.11	1.15	1.20	1.26	1.34	1.47	0.54^{***}	(16.65)
SIZE ($\times 10^{-9}$)	1.34	1.00	0.88	0.73	0.62	0.54	0.46	0.37	0.31	0.24	-1.11^{***}	(−6.75)
BM	0.51	0.53	0.53	0.54	0.54	0.53	0.53	0.53	0.50	0.45	-0.06^{**}	(−2.36)
MAX	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.05	0.03^{***}	(18.30)
TURN ($\times 10^{3}$)	4.12	4.27	4.48	4.72	4.86	5.08	5.29	5.59	5.90	7.31	3.19^{***}	(7.97)
STR ($\times 10^{3}$)	8.49	8.52	8.75	8.74	8.79	8.51	9.34	8.15	10.36	23.00	14.51^{***}	(4.28)

The table reports the time-series averages of the monthly cross-sectional median for stock characteristics of univariate decile portfolios formed based on MFD. Low (High) denotes the portfolio of stocks with the lowest (highest) MFD. The last two columns show the differences between the High and Low (H-L) and the associated Newey and West (1987) adjusted *t*-statistics (*t*-stat). The sample period is from August 1976 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

smaller, less liquid, and have higher idiosyncratic volatility and stronger lottery features.

The literature shows that the firm characteristics considered in Table 4 are useful in explaining the cross-section of expected stock returns. Stocks with higher asset growth, lower profitability, lower past 12-month (momentum) returns, lower earnings surprise, higher idiosyncratic volatility, and higher MAX tend to have lower future returns. Considering the prior findings in the literature and the fact that these firm characteristics vary across MFD deciles, it is important to control for the effects of investment, profitability, momentum, post-earnings-announcement drift, idiosyncratic volatility, and/or the lottery demand effect when studying the cross-sectional relation between MFD and future stock returns. Thus, in the next two subsections, we control for these well-known return predictors in bivariate portfolio sorts and in cross-sectional regressions to further test whether the significant relation between MFD and future stock returns remains intact.[12]

3.3 Bivariate portfolio-level analysis

Next, we investigate the negative association between MFD and future stock returns while controlling for the established equity return predictors. We conduct 5x10 dependent double sorts based on firm characteristics and MFD. Each month, we first sort stocks into quintile portfolios based on a given control. Then, we further sort stocks by MFD into decile portfolios within each control variable quintile. This bivariate portfolio analysis provides 50 conditionally double-sorted portfolios. Portfolio 1 (10) is the combined portfolio of stocks with the lowest (highest) MFD in each control variable quintile. Finally, we calculate the return spread between portfolio 10 and 1 for each control variable as well as its associated Fama and French (2018) six-factor alpha. We compute the return spreads for both equal- and value-weighted portfolios.

Table 5 presents the results. For brevity, we do not report the alphas for all 50 (5x10) portfolios. Instead, we report only the return spreads and alphas. Table 5 shows that the cross-sectional relation between MFD and future returns remains economically large and highly significant after controlling for a large set of well-known return predictors. The six-factor FF6 alpha spreads on the equal-weighted MFD-sorted portfolios are in the range of −0.64% per month (t-stat.=−6.35) and −0.89% per month (t-stat.=−7.73)

[12] In Table 3, we have already controlled for the market, size, value, momentum, investment, and profitability factors of Fama and French (2018) and Hou, Xue, and Zhang (2015) as well as the mispricing and behavioral factors of Stambaugh and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020) constructed based on earnings surprise (post-earnings-announcement drift) and a number of other well-known return predictors. As discussed in Section 4.1, the alpha spreads on MFD-sorted portfolios remain negative and highly significant in both value-weighted and equal-weighted portfolios after controlling for this large set of equity market factors.

Table 5
Bivariate portfolio sorts

	Equal-weighted		Value-weighted	
	H-L	FF6	H-L	FF6
SUE	−1.13***	−0.79***	−0.89***	−0.56***
	(−4.99)	(−6.81)	(−3.66)	(−3.44)
AG	−1.09***	−0.81***	−0.78***	−0.53***
	(−5.66)	(−7.23)	(−3.72)	(−3.52)
MOM	−1.07***	−0.80***	−0.90***	−0.67***
	(−5.31)	(−7.62)	(−4.31)	(−4.46)
ILLIQ	−1.29***	−0.88***	−1.06***	−0.63***
	(−5.60)	(−7.60)	(−4.56)	(−4.59)
OP	−1.02***	−0.86***	−0.76***	−0.54***
	(−6.42)	(−7.23)	(−3.53)	(−2.98)
IVOL	−0.95***	−0.76***	−0.91***	−0.69***
	(−7.76)	(−6.86)	(−5.55)	(−4.43)
BETA	−1.00***	−0.80***	−0.80***	−0.66***
	(−5.72)	(−6.96)	(−4.24)	(−4.10)
SIZE	−1.16***	−0.74***	−1.09***	−0.67***
	(−5.06)	(−6.08)	(−4.74)	(−5.56)
BM	−1.19***	−0.83***	−0.84***	−0.42***
	(−5.85)	(−7.31)	(−3.49)	(−2.88)
MAX	−0.98***	−0.64***	−0.76***	−0.46***
	(−6.87)	(−6.35)	(−4.49)	(−2.94)
TURN	−1.22***	−0.83***	−0.85***	−0.56***
	(−6.17)	(−7.09)	(−4.19)	(−3.35)
STR	−1.27***	−0.89***	−1.01***	−0.64***
	(−6.30)	(−7.73)	(−4.48)	(−3.79)

The table reports results from bivariate portfolios based on dependent double sorts of various firm-specific characteristics and MFD. First, quintile portfolios are formed every month based on a firm-specific characteristic. Next, additional decile portfolios are formed based on MFD within each firm-specific characteristic quintile. Subsequently, we average returns for each MFD decile across the characteristic quintiles, yielding 10 quintile-mean decile returns. Finally, we report the difference-in-differences return spreads between the lowest and highest MFD decile returns as well as the associated Fama and French (2018) six-factor alpha. We consider equal- and value-weighted portfolios. The stock characteristics are described in Table 1. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from August 1976 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

and ranging from −0.42% per month (t-stat.=−2.88) to −0.69% per month (t-stat.=−4.43) for value-weighted bivariate sorts. These results indicate that even after controlling for various firm characteristics and risk factors in bivariate portfolios, there is a strong negative relation between MFD and future equity returns. In other words, the predictive power of MFD is not explained by other cross-sectional return predictors, including the existing measures of investor disagreement.

3.4 Fama-MacBeth cross-sectional regressions

In this section, we conduct firm-level Fama-MacBeth regression analysis to test whether MFD predicts the cross-section of future stock returns while controlling for other known predictors simultaneously. Each month, we run a cross-sectional regression of stock returns in that month on past MFD, as well as a number of control variables, including the 1-month-lagged market beta, size, book-to-market, momentum, operating profitability, asset growth, earnings surprise, short-term return reversal, illiquidity, turnover ratio, idiosyncratic volatility, and lottery demand. The stock-level cross-sectional regressions are run each month and the standard errors of the average slope coefficients are corrected for heteroscedasticity and autocorrelation following Newey and West (1987).

Table 6 reports the results of stock-level Fama-MacBeth regressions. In column 1, we include MFD as well as beta, size, book-to-market, and momentum as additional cross-sectional predictors. Consistent with the portfolio results, we find a negative and significant relation between MFD and 1-month-ahead returns. The average slope coefficient on MFD is −0.35 with a t-statistic of −7.11. In columns 2 and 3 we include additional return predictors in the cross-sectional regressions. Even in the presence of 12 well-known predictors, the average slope coefficient is −0.24 and statistically significant with a t-statistic of −5.35. MFD is also highly economically significant. The spread in the average standardized MFD between deciles 10 and 1 is approximately 3.39, and multiplying this spread by the average slope of −0.24 yields a return difference of −0.81% per month, controlling for all else. In most cases, the slope coefficients on the control variables are consistent with prior literature; short term reversal (STR), turnover (TURN), asset growth (AG), and MAX are negatively correlated with the future return, whereas momentum (MOM), profitability (OP), and earnings surprise (SUE) are positively related to the next month's return.

In column 4, we include the industry-adjusted return in month $t+1$ to account for the industry effect. Specifically, we adjust the dependent variable by subtracting the firm's value-weighted Fama-French 48-industry return from the firm's current month return. Doing so allows us to tease out the return predictive power from MFD rather than the 1-month industry momentum effect (Moskowitz and Grinblatt 1999). The coefficient of MFD remains similar after controlling for the industry return directly. In column 5, we further control for the common characteristics that are shown to affect stock returns systematically. Specifically, we follow Daniel et al. (1997) and compute the characteristics-adjusted returns as the difference between the firm's return and the corresponding DGTW benchmark portfolio returns. We replace the firm's raw return with this characteristics-adjusted return as the dependent variable and run the same set of monthly cross-sectional regressions. Again, the magnitude of the slope coefficient on MFD becomes slightly weaker, but it remains highly significant, both economically and statistically.

4. Robustness Checks

The strong negative association between MFD and future returns is robust to a comprehensive battery of tests, which we describe in the Internet Appendix. We confirm that the predictive power of MFD is

Table 6
Fama-MacBeth cross-sectional regressions

	Excess return	Excess return	Excess return	Industry-adj. return	DGTW-adj. return
Const	0.86***	0.82***	0.84***	0.08	0.03
	(3.64)	(3.44)	(3.58)	(0.80)	(0.77)
MFD	−0.35***	−0.31***	−0.24***	−0.21***	−0.19***
	(−7.11)	(−7.45)	(−5.35)	(−5.84)	(−4.72)
BETA	0.02	0.04	0.07	0.09**	0.06
	(0.29)	(0.55)	(1.27)	(2.41)	(1.37)
SIZE	−0.04**	−0.05**	−0.06***	−0.04***	−0.04***
	(−1.98)	(−2.13)	(−2.69)	(−2.79)	(−3.03)
BM	0.13**	0.14**	0.08	0.09**	−0.03
	(2.39)	(2.57)	(1.48)	(2.25)	(−0.82)
MOM	0.38***	0.38***	0.37***	0.31***	0.23***
	(6.50)	(6.69)	(5.84)	(5.90)	(4.76)
AG		−0.26***	−0.24***	−0.19***	−0.19***
		(−5.00)	(−4.13)	(−4.18)	(−4.02)
OP		0.16***	0.14***	0.14***	0.15***
		(3.59)	(3.29)	(3.24)	(3.28)
SUE			0.07***	0.06***	0.05***
			(3.01)	(3.40)	(2.58)
ILLIQ			0.02	0.09	0.04
			(0.76)	(1.40)	(0.62)
IVOL			0.01	−0.04	0.02
			(0.08)	(−0.62)	(0.41)
MAX			−0.16***	−0.12**	−0.14**
			(−2.91)	(−2.40)	(−2.46)
TURN			−0.15***	−0.12***	−0.12***
			(−3.51)	(−3.00)	(−2.88)
STR			−0.28***	−0.36***	−0.33***
			(−4.88)	(−6.61)	(−5.58)
Observations	1,167,280	1,102,119	978,110	934,973	934,973

The table reports Fama-MacBeth cross-sectional regressions for MFD. MFD and the control variables in month $t-1$ are matched to stock returns in month t. The dependent variable is the firm's future excess return in the first three columns (Excess return), the firm's future return over its value-weighted industry peers' return (Industry-adj. return), or the firm's DGTW adjusted return (DGTW-adj. return). All dependent variables are given in percent. The control variables are described in Table 1, winsorized at 0.5% in both tails, and standardized. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from August 1976 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

persistent, lasting for several months after portfolio formation (see Internet Appendix A). The MFD premium is not subsumed by other known anomalies in extensive bivariate sorts (Internet Appendix B) and remains highly significant when estimated using the three-pass methodology of Giglio and Xiu (2021), which is robust to omitted priced factors (Internet Appendix C). The MFD premium generates large and statistically significant pricing errors when evaluated against recently developed nonlinear stochastic discount factors (e.g., Chen, Pelger, and Zhu 2024; Cong et al. 2025), confirming that it captures mispricing unexplained by complex, nonlinear risks (Internet Appendix D). Furthermore, our findings are not sensitive to the specific machine learning architecture or hyperparameter choices used to generate belief dispersion; the results hold when using alternative models like gradient boosted trees or penalized linear models (Internet Appendix E). To ensure MFD is not simply a proxy for stocks that are inherently difficult to forecast, we confirm its predictive power holds strongly even among the easiest-to-predict stocks (see Internet Appendix F). Additionally, the MFD premium holds internationally across developed and emerging markets, providing strong evidence of external validity (Internet Appendix G).

Our methodology also allows us to distinguish between different economic sources of investor disagreement, which maps naturally onto the fundamental asset pricing decomposition of expected returns into discount rates and cash flows. While disagreement over future returns (our main MFD measure) can be interpreted as heterogeneity in beliefs about discount rates, disagreement over future earnings relates to heterogeneity in beliefs about cash flows. We test this latter channel by constructing an alternative MFD measure based on forecasts of future earnings. As detailed in Internet Appendix H, we find compelling evidence that this earnings-based MFD is also a strong negative predictor of future returns, highlighting the importance of disagreement about firms' fundamental cash flows in driving the MFD premium. To formally assess the relative importance of the discount rate versus the cash flow channel, we construct a joint MFD index using decile ranks of both return-based and one of the three earnings-based measures. Regressing future returns on this joint index confirms its strong negative predictive power. We then decompose the effect using the methodology of Hou and Loh (2016). The results show that both channels are highly significant drivers of the joint effect, while disagreement about cash flows accounts for a slightly larger portion of the joint effect.

5. MFD as a Measure of Disagreement

In this section, we provide evidence that MFD captures investor disagreement. We first compare it specifically to AFD in Section 5.1 before providing evidence from additional disagreement proxies in Section 5.2. We decompose the effect of a disagreement index on future excess returns with respect to MFD in Section 5.3. In Section 5.4, we focus on the cross-sectional association between MFD and future stock returns in different market phases. We study the predictive power of MFD for trading volume and volatility in Section 5.5. We decompose the total disagreement effect into components related to model and information set disagreement, respectively, in Section 5.6.

5.1 Comparing MFD to analyst-based disagreement

In this section, we benchmark MFD against analyst forecast dispersion (AFD) in two stages. First, we establish that MFD and AFD capture a

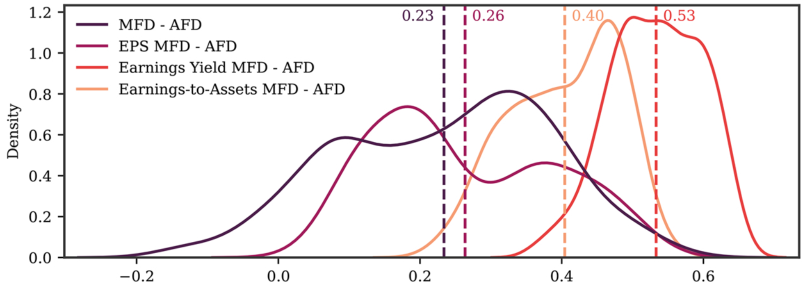

Figure 2
Cross-sectional rank correlation between AFD and MFD
The figure shows the density of the monthly cross-sectional correlation between analyst forecast dispersion (AFD, Diether, Malloy, and Scherbina 2002) and MFD based on future excess returns and realized earnings. Each month t, the correlation between MFD and AFD is measured using Spearman's ρ. Besides return-based MFD, the figure shows the cross-sectional rank correlations between three earnings-based MFD measures and AFD: MFD based on earnings-per-share (EPS MFD), earnings-yield (Earnings yield MFD), and earnings-to-assets (Earnings-to-asset MFD). The vertical dashed red lines represent the average monthly cross-sectional rank correlations. The sample period is from January 1983 to March 2022 for return-based MFD and from February 1994 to March 2022 for the earnings-based MFD measures, respectively.

common underlying economic phenomenon by examining their statistical correlation and shared economic drivers. Second, we demonstrate that MFD is a more powerful and robust predictor of future stock returns.

We begin by documenting a significant positive relationship between MFD and AFD. Figure 2 shows the distribution of the monthly cross-sectional rank correlations between MFD and AFD based on Spearman's ρ. The average monthly cross-sectional correlation is 0.23 and positive in almost all months. Figure 2 also presents the cross-sectional rank correlation between earnings-based MFD and AFD. The cross-sectional correlations between AFD and our earnings-based measures of MFD are higher; 0.26 for earnings-per-share MFD, 0.40 for earnings-to-asset MFD, and 0.53 for earnings yield MFD. Traditional analyst forecasts, and hence AFD, are centered on future earnings, which are a direct proxy for future cash flows. Our earnings-based MFD measures are constructed similarly and thus capture disagreement about the cash flow component of valuation. In contrast, our primary return-based MFD measure captures disagreement about the total expected return, which incorporates both cash flow and discount rate expectations. The stronger correlation between earnings-based MFD and AFD suggests both are primarily capturing expected cash flow disagreement. The comparatively lower correlation between our return-based MFD and AFD highlights that our main measure captures an additional, distinct dimension of disagreement related to discount rates. However, even the comparatively low correlation between return-based MFD and AFD is remarkable in light of the findings of Goulding, Harvey, and Kurtović (2025). The authors

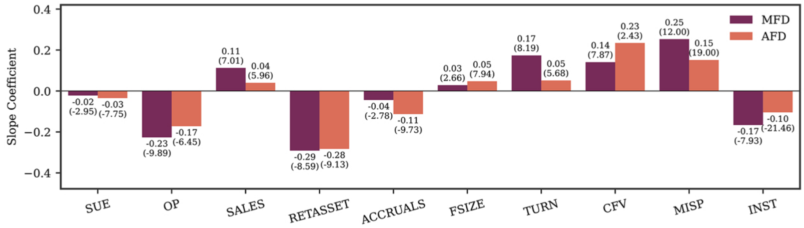

Figure 3
Cross-sectional determinants of MFD and AFD
This figure displays the time-series average of slope coefficients from monthly Fama-MacBeth regressions of MFD and AFD on various firm characteristics. We run parallel Fama-MacBeth regressions of AFD and MFD on a set of core controls (firm size, book-to-market value, and historical stock beta over the last 60 months) and firm characteristics associated with information complexity and uncertainty one-by-one. The latter characteristics comprise standardized unexpected earnings (SUE), profitability (OP), sales growth (SALES), return on assets (RETASSET), accruals (ACCRUALS), the log of the net file size of SEC 10-K filings (FSIZE, obtained via https://sraf.nd.edu/complexity/), turnover (TURN), cash-flow volatility (CFV), the stock mispricing (MISP) index of Stambaugh, Yu, and Yuan (2015), and institutional ownership (INST). The bars are annotated with the respective slope coefficient and Newey-West t-statistics are reported in parentheses. The sample period is from August 1976 (January 1983) to December (March) 2022 for MFD (AFD).

compare the correlations among major investor disagreement measures and find that the average correlation among these measures is below 0.15, which is exceeded in all our cases.

To further show that MFD and AFD capture a common underlying economic phenomenon, we conduct a detailed investigation into their economic drivers. We run parallel Fama-MacBeth regressions of MFD and AFD on a core set of controls (stock market beta, firm size, and firms' book-to-market value) and add firm characteristics associated with information complexity and uncertainty one-by-one. The results, reported in Figure 3, show that MFD and AFD are not just statistically correlated but are rooted in the same fundamental firm-level conditions. Both MFD and AFD are significantly higher for firms with lower profitability, lower return-on-assets, higher cash flow volatility, and greater stock mispricing scores according to Stambaugh, Yu, and Yuan (2015). Furthermore, both measures increase with firm complexity, as proxied by the log net file size of SEC filings (Loughran and McDonald 2014), and are higher for firms with lower institutional ownership proxying for more binding short-sale constraints (Nagel 2005; Sikorskaya 2024). Further analysis, presented in Figure I.1 in the Internet Appendix, shows their long-short strategy premiums also share common time-series drivers, becoming most pronounced during periods of high aggregate uncertainty and sentiment.

We provide further evidence on the positive correlation between AFD and MFD using portfolio sorts. Panel A of Table 7 depicts the average AFD per MFD decile in univariate portfolio sorts on MFD. AFD is monotonically increasing in MFD deciles and the spread in AFD between MFD deciles 10 and 1 is 0.18 with a t-statistic of 17.73. Additional time-series evidence

Table 7
Analyst forecast dispersion and MFD

A. Average AFD in MFD decile portfolio

	Low	2	3	4	5	6	7	8	9	High	H-L	t-stat
AFD	0.08	0.10	0.11	0.13	0.15	0.16	0.18	0.20	0.22	0.25	0.18***	17.73

B: Bivariate Portfolio Sort on AFD

	Low	2	3	4	5	6	7	8	9	High	H-L	t-stat	FF6	t-stat
AFD low	1.28	1.06	1.20	1.13	1.14	0.99	0.99	0.90	0.86	0.62	−0.66***	−2.71	−0.40**	−2.32
AFD 2	0.93	1.13	1.01	0.85	1.01	0.92	0.79	0.66	0.55	0.13	−0.80***	−3.17	−0.46**	−2.41
AFD 3	1.08	0.90	1.10	0.96	0.80	0.86	0.71	0.84	0.76	0.04	−1.03***	−3.72	−0.64***	−3.20
AFD 4	0.82	1.07	0.91	0.98	0.77	0.71	0.70	0.69	0.42	−0.05	−0.87***	−2.59	−0.35	−1.43
AFD high	0.79	0.77	0.80	0.49	0.43	0.48	0.49	0.49	0.20	−0.51	−1.30***	−3.70	−0.94***	−3.29
AFD H-L	−0.49	−0.29	−0.40	−0.64	−0.71	−0.51	−0.51	−0.41	−0.65	−1.13	−0.64**	−2.16	−0.54*	−1.75

Panel A reports the average analyst forecast dispersion (AFD) of the MFD-sorted univariate decile portfolios. Low (high) AFD indicates a lower (higher) average forecast dispersion. Panel B reports 5x10 dependent bivariate equal-weighted portfolio sorts. First, quintile portfolios are formed every month using AFD. Next, decile portfolios are formed based on MFD within each firm-specific AFD quintile. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from January 1983 to December 2022.
* p <.1; ** p <.05; *** p <.01.

on the overlap between AFD and MFD is presented in Figure I.2 in the Internet Appendix. Panel A of Figure I.2 shows the yearly average median AFD per MFD decile rank portfolios, whereas panel B depicts the median AFD rank per MFD decile rank. We focus on ranks 1, 5, and 10 as we are mainly interested in the overlap in extreme deciles. We also include earnings-yield MFD as it has shown the highest correlation to AFD in Figure 2. A clear monotonic relationship exists across three MFD rank portfolios (1, 5, 10), with higher MFD ranks consistently corresponding to higher AFD values and ranks throughout the sample period. The highest MFD rank portfolio shows significantly higher AFD values and ranks compared to the lowest MFD rank portfolio. Similar to Figure 2, AFD shows generally a higher overlap to earnings yield MFD. This gap is more pronounced for AFD values instead of ranks. For the latter, panel B shows that the gap between earnings-yield MFD and standard MFD is less pronounced in terms of rank ordering.

In the second stage, we investigate the strength of the cross-sectional predictions from MFD and AFD. We begin with bivariate portfolio analysis in which we first sort stocks into quintile portfolios every month based on AFD. Subsequently, we divide each AFD quintile into deciles based on MFD. Panel B of Table 7 reports the bivariate portfolio results. The MFD decile return spread is statistically significant in all AFD quintiles. The MFD return spread becomes larger in magnitude and more significant when analyst disagreement is more severe; in AFD quintile five, the equal-weighted MFD return spread is −1.30% (t-stat.=−3.70). Moreover, the corresponding FF6 alpha is statistically significant in all AFD quintiles, except for the fourth AFD quintile. The difference in MFD high-minus-low return spreads between the highest and lowest AFD quintile is economically large and statistically significant, similarly for the FF6 alpha spread.

Finally, prior evidence on the cross-sectional association between AFD and stock returns is mixed.[13] We revisit the evidence on AFD using our longer time period. We analyze AFD-sorted portfolios in the same way we did for MFD. Table I.1 in the Internet Appendix shows the equal-weighted and value-weighted decile portfolio returns as well as the return and alpha spreads between high-AFD and low-AFD decile portfolios. The Fama and French (2018) six-factor alpha spread is −0.60% with a t-statistic of −4.79

[13] Diether, Malloy, and Scherbina (2002), Chen, Hong, and Stein (2002), Goetzmann and Massa (2005), Berkman et al. (2009), and Yu (2011) find a negative cross-sectional association between AFD and average stock returns. Others present evidence that the negative relation holds only for a sample of stocks with certain characteristics, for example, small, illiquid, low credit quality, or short sale constrained. In particular, Malkiel and Cragg (1970), Qu, Starks, and Yan (2003), Doukas, Kim, and Pantzalis (2006), Avramov et al. (2009), and Carlin, Longstaff, and Matoba (2014) find either a positive or no significant relation between AFD and future stock returns.

for the equal-weighted portfolios, whereas the alpha spread is much lower at −0.15% and statistically insignificant for the value-weighted portfolios. The evidence for AFD aligns with our findings for MFD, but the effect is notably weaker in both magnitude and statistical significance. Finally, we compare the relative performance of MFD and AFD using stock-level Fama-MacBeth regressions controlling for other well-known return predictors. Table I.2 in the Internet Appendix shows that the inclusion of AFD does not influence the statistical and economic significance of MFD. More importantly, MFD exhibits a nearly 50% to 70% larger effect in economic terms compared to AFD, even after controlling for other return predictors. Collectively, these results stress the relatively higher predictive power of MFD with respect to AFD and shows that the MFD effect is much stronger for equities with high dispersion in analysts' earnings forecasts.

5.2 Correlation to additional disagreement proxies

To further validate MFD, we examine its alignment with a broad range of disagreement proxies used in the prior literature. These include measures based on volatility (idiosyncratic volatility), trading activity (turnover and unexplained volume), social media sentiment (StockTwits disagreement), option market positioning, and analyst coverage patterns (see Internet Appendix I.3 for detailed descriptions).

We calculate the time series of cross-sectional rank correlations using Spearman's ρ between the additional disagreement proxies and MFD. We also compute the cross-sectional correlations of the aforementioned proxies to AFD. Table 8 presents the results. We document strong average rank correlations between MFD and all additional disagreement proxies. All average rank correlations are statistically significant. Additionally, the rank correlations to AFD are weaker for all disagreement proxies and the difference is statistically significant in all cases, except for new analyst issues and expected idiosyncratic skewness. For the latter, the difference between the rank correlations is not statistically significantly different from zero. Overall, Table 8 indicates that MFD better aligns with previously proposed measures of investor disagreement.

5.3 Decomposing a disagreement index

To formally test MFD's ability to explain the return premium associated with broad-based stock-level disagreement, we construct a composite index from several non-analyst-based proxies, following Goulding, Harvey, and Kurtović (2025), and decompose its predictive power. The index averages the decile ranks of four proxies: historical volatility, short-interest, idiosyncratic volatility, and expected idiosyncratic skewness. We do not include analyst-based disagreement measures as we want to compare and dissect the aggregate disagreement index using MFD and AFD.

Table 8
Average cross-sectional rank correlations to MFD and AFD for disagreement proxies

	MFD		AFD		Difference	
	XS-Corr.	t-stat	XS-Corr.	t-stat	XS-Corr.	t-stat
HV	0.45***	(27.51)	0.29***	(24.63)	0.16***	(9.64)
IVOL	0.49***	(23.74)	0.37***	(28.18)	0.11***	(5.84)
StockTwits disagreement	0.16***	(13.62)	0.08***	(6.62)	0.08***	(7.64)
StockTwits within group disagreement	0.15***	(18.25)	0.06***	(9.16)	0.09***	(11.86)
StockTwits across group disagreement	0.09***	(8.99)	0.04***	(4.41)	0.05***	(5.95)
Last month turnover	0.16***	(12.87)	0.11***	(12.27)	0.05***	(3.46)
Standardized unexplained volume	0.07***	(10.64)	0.00	(0.02)	0.07***	(14.10)
Option disagreement	0.03**	(2.07)	–0.03***	(−4.26)	0.05***	(5.04)
Expected idiosyncratic volatility	0.21***	(9.94)	0.21***	(13.98)	-0.00	(−0.37)
New analyst issues	0.03***	(7.02)	0.03***	(10.91)	-0.00	(−0.96)

The table reports the average cross-sectional rank correlations between MFD and AFD and other commonly used disagreement proxies. Other disagreement proxies comprise idiosyncratic volatility (Boehme, Danielsen, and Sorescu 2006; Berkman et al. 2009), Stockwits disagreement (Cookson and Niessner 2020, 2023), and monthly turnover. Idiosyncratic volatility is measured as the standard deviation of the daily residuals estimated from the regression of the daily excess stock returns on the daily market return over the previous year. StockTwits disagreement is calculated in three ways (Cookson and Niessner 2020, 2023): overall, within and across investment approaches. Standardized unexplained volume is constructed following Garfinkel (2009). Option disagreement follows Ge, Lin, and Pearson (2016) and Golez and Goyenko (2022). Expected idiosyncratic volatility is taken from Boyer, Mitton, and Vorkink (2010). New analyst issues follows the logic in Goulding, Harvey, and Kurtović (2025). Details are given in Section I.3 in the Internet Appendix. Newey and West (1987) adjusted $t-$statistics are reported in parentheses. The last column (Difference) shows the average differences and their t-statistics in cross-sectional rank correlations between MFD and disagreement proxies versus AFD and disagreement proxies. The sample period is from August 1976 to December 2022 for idiosyncratic volatility, monthly turnover, standardized unexplained volume and expected idiosyncratic volatility, whereas it ranges from January 2010 to December 2021 for StockTwits data. The sample ranges from May 2005 to February 2021 for option disagreement. The sample for new analyst issues ranges from January 1983 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

We first study the overlap of the disagreement index with MFD and AFD, respectively. For the AFD sample, we show the median disagreement index for MFD- and AFD-sorted portfolios in Figure 4. We again focus on ranks 1, 5, and 10. The figure shows a consistent hierarchical pattern. The highest MFD decile exhibits significantly higher median disagreement ranks (8–10 range) than deciles 5 and 1 (declining from 5 to 2–3 over time in case of the latter). The alignment of the disagreement index with MFD appears to become stronger over time as the highest (lowest) decile shows a clear upward (downward) trend over the sample period. Notably, the alignment between the disagreement index and AFD is consistently weaker. The average disagreement index rank is up to two ranks lower for the highest AFD decile compared to MFD. It is similar for the lowest AFD decile rank.

Next, we proceed to the decomposition analysis. After confirming that the disagreement index is a strong negative return predictor (as shown in the first column of Table 9), we employ the methodology of Hou

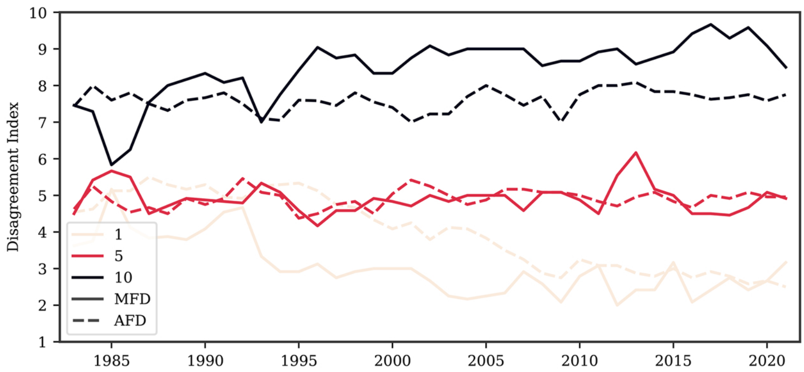

Figure 4
Median disagreement index rank per MFD and AFD ranks
The figure shows the overlap of a disagreement index at the stock level with MFD and AFD, respectively. We construct a disagreement index at the stock level using historical volatility, short-interest, idiosyncratic volatility and expected idiosyncratic skewness. Each month, we classify cross-sectionally stocks into 10 ranks for each component. Thereby a higher rank indicates higher disagreement. Subsequently, the disagreement index is the average of the ranks. The figure shows the yearly median disagreement index rank per MFD and AFD ranks of 1, 5, and 10. Solid lines represent disagreement index ranks per MFD ranks, whereas dashed lines represent disagreement index ranks per AFD ranks. The sample period is from January 1983 to December 2022.

and Loh (2016) to attribute this effect to MFD and AFD.[14] The results are presented in Table 9. We consider two sets of stocks samples and portfolios. The first uses our entire MFD sample, whereas the latter is confined to the availability of AFD. In univariate decompositions at the stock level, depicted in panel A in Table 9, MFD explains over 50% of the index's effect, comparing well with AFD explaining only 11%. This dominance is stronger in a multivariate decomposition that simultaneously considers both measures: MFD's explanatory power remains high at over 54%, whereas the fraction attributable to AFD falls to 7%. The findings are even stronger in portfolio-level tests designed to mitigate measurement error, where MFD explains nearly 83% of the disagreement index effect. Hence, these results establish that MFD captures the lion's share of the return-relevant information contained in a broad set of disagreement proxies.

5.4 Evidence from different market phases

A potential concern is that MFD's predictive power merely reflects time variation in the market's cross-sectional factor structure. In periods

[14] Hou and Loh (2016) introduce a regression methodology to evaluate a number of potential explanations for the puzzling negative relation between idiosyncratic volatility and future stock returns. They find that the idiosyncratic volatility puzzle is largely driven by investors' preferences for lottery-like stocks and market frictions. In this section, we implement the decomposition methodology proposed by Hou and Loh (2016) to examine whether MFD or AFD explains a larger fraction of the disagreement effect.

Table 9
Decomposition of disagreement index (DIS) on future excess returns with respect to MFD and AFD

A: Single stocks

	DIS on Excess returns			Decomposition wrt MFD			Decomposition wrt AFD			Multivariate decomposition		
MFD sample	Const.	1.55***	(8.39)	MFD	50.84%***	(6.75)						
	DIS	−0.13***	(−3.33)	Residual	49.16%***	(6.53)						
AFD sample	Const.	1.39***	(7.51)	MFD	56.34%***	(5.26)	AFD	11.42%***	(3.77)	MFD	54.48%***	(5.23)
	DIS	−0.12***	(−2.99)	Residual	43.66%***	(4.08)	Residual	88.58%***	(29.23)	AFD	7.45%***	(3.46)
										Residual	38.07%***	(3.34)

B: 50 portfolios sorted on DIS

	DIS on Excess returns			Decomposition wrt MFD			Decomposition wrt AFD			Multivariate decomposition		
MFD sample	Const.	1.60***	(8.40)	MFD	82.80%***	(12.18)						
	DIS	-0.15***	(-3.62)	Residual	17.20%**	(2.53)						
AFD sample	Const.	1.45***	(7.99)	MFD	90.10%***	(13.96)	AFD	64.91%***	(9.80)	MFD	69.62%***	(10.41)
	DIS	-0.13***	(-3.46)	Residual	9.90%	(1.53)	Residual	35.09%***	(5.30)	AFD	26.45%***	(5.06)
										Residual	3.93%	(0.81)

The table shows the decomposition of a disagreement index on future excess returns with respect to MFD and AFD. We employ the decomposition methodology of Hou and Loh (2016). We first construct a disagreement index (DIS) at the stock level using historical volatility computed from daily returns over the last month, short-interest, idiosyncratic volatility computed using the CAPM and daily returns over the last year, and expected idiosyncratic skewness. The first column (DIS on excess returns) presents results from regressing future stock returns cross-sectionally on the disagreement index (the first stage according to Hou and Loh 2016). The second column (Decomposition wrt MFD) univariately decomposes the effect of the disagreement index on future stock returns with respect to MFD (the second stage according to Hou and Loh 2016). The third column (Decomposition wrt AFD) applies the univariate decomposition with respect to AFD. The last column (Multivariate decomposition) applies a multivariate decomposition on the effect of the disagreement index on future stock returns with respect to MFD and AFD. The table reports the fraction of the effect of the disagreement index on future stock returns which can be explained by MFD and AFD, respectively. It also presents the residual fraction which cannot be explained by MFD and AFD, respectively. The decomposition with respect to MFD is separately shown for the larger MFD sample. Panel A reports results for single stocks, whereas panel B uses 50 portfolios sorted on the disagreement index as assets in the cross-sectional regressions. t-statistics are given in parentheses. The sample period is from January 1978 to December 2021 for the MFD sample (because of the availability of expected idiosyncratic skewness), whereas it ranges from January 1983 to December 2021 for the AFD sample. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Table 10
Performance of value-weight MFD univariate long-short portfolio in different market phases

	High		Low		Difference	
	Return	t-stat	Return	*t*-stat	Return	*t*-stat
Factor density	−1.48***	(−3.48)	−0.64**	(−2.45)	−0.84**	(−1.73)
VIX Index	−1.84***	(−3.20)	−0.18	(−0.57)	−1.67***	(−2.56)
Volatility-of-volatility	−1.76***	(−3.09)	−0.26	(−0.66)	−1.50**	(−2.30)
VVIX Index	−1.29**	(−2.07)	0.22	(0.37)	−1.51**	(−1.84)
CATFIN	−2.38***	(−5.34)	0.18	(0.62)	−2.56***	(−5.35)
Financial uncertainty	−1.48***	(−3.48)	−0.70***	(−2.62)	−0.79*	(−1.60)
Real uncertainty	−1.35***	(−3.34)	−0.83***	(−2.66)	−0.52	(−1.07)
Macro uncertaingy	−1.49***	(−3.70)	−0.69**	(−2.54)	−0.80*	(−1.63)
PLS Sentiment Index	−1.89***	(−4.60)	−0.30	(−1.10)	−1.59***	(−3.27)
Baker-Wurgler (orthogonalized)	−1.83***	(−5.08)	−0.24	(−0.76)	−1.60***	(−3.32)
Aggregate IVOL	−1.60***	(−3.68)	−0.58**	(−2.28)	−1.02**	(−2.08)
Standardized unexplained stock volume	−1.45***	(−4.02)	−0.73**	(−2.44)	−0.71*	(−1.46)
Option disagreement	−1.83***	(−3.01)	−0.01	(−0.03)	−1.82***	(−2.73)

The table reports returns to the value-weight MFD long-short portfolio for different sample splits capturing states of factor density and sparsity, and high and low uncertainty and disagreement. For factor density, we use 120-month rolling windows to estimate the number of nonzero stock characteristics applying Lasso to the cross-section of monthly stock returns. 153 stock characteristics are taken from Jensen, Kelly, and Pedersen (2023). The entire sample is split by the median of the nonzero stock characteristics. Furthermore, our sample is split by the median VIX Index from January 1990 to December 2022, the median of volatility-of-volatility from January 1990 to December 2022, the VVIX Index from March 2006 to December 2022, the median of the aggregate systemic risk index (CATFIN) of Allen, Bali, and Tang (2012) from August 1976 to December 2022, the median of the financial, real, and macro uncertainty indices of Jurado, Ludvigson, and Ng (2015) from August 1976 to December 2022, the sentiment index (PLS Sentiment Index) of Huang et al. (2015) from August 1976 to December 2022, the sentiment index (Baker-Wurgler [orthogonalized]) of Baker and Wurgler (2007) orthogonalized to macroeconomic uncertainty from August 1976 to December 2022, aggregate idiosyncratic volatility (Aggregate IVOL) from August 1976 to December 2022, standardized unexplained stock volume from August 1976 to December 2022, and option disagreement from January 1990 to May 2020. Details are given in Internet Appendix I.3. The last columns (Difference) gives shows the difference in mean realized returns between high and low uncertainty/disagreement states as well as its statistical significance using a one-sided *t*-test. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

when returns are described by a few dominant factors (sparsity), model forecasts should converge, leading to low MFD. Conversely, when the factor structure is dense, forecasts may diverge, elevating MFD. To test this hypothesis, we classify periods as sparse or dense based on the time-series median of the number of predictive characteristics selected by a rolling 10-year window LASSO regression using all 153 characteristics. Finally, we measure the value-weighted return spread between MFD-sorted decile portfolios 10 and 1 across both subsamples. The first row in Table 10 reports the results. Factor density attenuates the magnitude of the MFD return. However, MFD still yields substantial and statistically significant returns of approximately 7.68% per annum in times of sparsity. This finding suggests that our MFD measure captures a dimension of disagreement beyond what can be explained by time variation in standard factor exposures alone.

Motivated by the above subsample analysis, we conclude this section by evaluating the cross-sectional association of MFD with future stock returns across different market phases. If MFD captures disagreement, its value-weighted return spread should be more pronounced in absolute magnitude during times of higher disagreement and uncertainty. We classify months into high and low periods of uncertainty/disagreement by sample splits based on the median of aggregate uncertainty and disagreement proxies. We consider the following proxies: the VIX index, volatility of aggregate volatility (Agarwal, Arisoy, and Naik 2017), the aggregate systemic risk index of Allen, Bali, and Tang (2012), the financial, real, and macro uncertainty indices of Jurado, Ludvigson, and Ng (2015), the sentiment index of Huang et al. (2015), the sentiment index orthogonalized to macroeconomic shocks of Baker and Wurgler (2007), aggregate idiosyncratic volatility, standardized unexplained stock volume, and disagreement in the S&P 500 index option market.[15]

The results, reported in Table 10, strongly support this prediction. During periods of high uncertainty and disagreement, the MFD return spread is economically large and statistically significant across nearly all proxies, with monthly returns ranging from −1.29% to −2.38%, the latter being more than double the unconditional average presented in Table 3. Conversely, the premium is substantially attenuated during periods of low disagreement and is often statistically indistinguishable from zero. Furthermore, the difference in the MFD premium between high and low states is itself statistically significant for almost all proxies tested using a one-sided t−test. This state-dependent behavior provides further evidence that MFD captures investor disagreement.

5.5 Explanatory power for trading volume and volatility

The previous literature has shown that trading volume is increasing in belief disagreement (Bessembinder, Chan, and Seguin 1996; Goetzmann and Massa 2005; Cookson and Niessner 2020). Atmaz and Basak (2018) build a model of belief disagreement with a continuum of investors differing in beliefs that matches the empirical observation. If dispersion is higher, investors with more divergent beliefs and consequently greater trading needs play a more significant role. Additionally, there also exists a positive empirical and theoretical link between disagreement and volatility (Scheinkman and Xiong 2003; Banerjee 2011; Atmaz and Basak 2018). Based on these insights, we study more closely the relation between trading volume and volatility by means of panel regressions. Precisely, we estimate the following panel regressions:

$$y_{i,t} = \alpha_i + \gamma_t + \beta_1 \times MFD_{i,t} + \beta_2 \times y_{i,t-1} + \gamma \times Controls_{i,t} + \epsilon_{i,t}, \tag{5}$$

where $y_{i,t}$ is standardized unexplained volume, monthly turnover, or the historical volatility of stock i in month t. Our coefficient of interest is β_1. We also include the first lag of our dependent variable as a right-hand

[15] We detail the construction of aggregate uncertainty and disagreement proxies in Section I.4 of the Internet Appendix for proxies we construct ourselves as public data are not available.

Table 11
Predicting standardized unexplained volume, monthly turnover and historical volatility

	SUV		Turnover		Historical volatility	
	t	t+1	t	t+1	t	t+1
MFD	0.08***	0.01***	0.11***	0.01***	0.26***	0.04***
	(25.92)	(3.31)	(29.15)	(5.32)	(34.05)	(13.99)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Entity FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,095,384	1,091,603	1,167,263	1,162,122	1,065,430	1,059,353
R-squared	.08	.08	.38	.36	.21	.11

The table reports results for panel regressions of trading volume and stock volatility on MFD. For trading volume, we consider standardized unexplained volume (SUV) and monthly turnover. For volatility, we use the standard deviation of daily stock returns in a month. We estimate the following regression $y_{i,t} = \alpha_i + \gamma_t + \beta_1 \times MFD_{i,t} + \beta_2 \times y_{i,t-1} + \gamma \times Controls_{i,t} + \epsilon_{i,t}$, where $y_{i,t}$ is either standardized unexplained volume, monthly turnover, or the historical volatility of stock i in month t. We also include the first lag of our dependent variable to account for persistence in volume measures and volatility, respectively, and add day (γ_t) and stock fixed effects (α_i). Additional controls ($Controls_{i,t}$) consist of the book-to-market ratio, stock beta, market capitalization, 12-1 month momentum and short-term reversal. Standard errors are double-clustered by month and firm. The dependent and independent variables are standardized and winsorized at 0.5% in both tails. The sample period is from January 1978 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

side variable to account for persistence in volume measures and volatility, respectively, and add day (γ_t) and stock fixed effects (α_i). Additional controls ($Controls_{i,t}$) consist of the book-to-market ratio, stock beta, market capitalization, 12-1 month momentum and short-term reversal. Standard errors are double-clustered by month and firm.

Table 11 presents our results on the link between MFD and measures of trading volume and volatility. We observe a strong relation between MFD and trading volume and historical volatility in economic and statistical terms. The coefficient on MFD is 0.08, 0.11, and 0.26 for standardized unexplained volume, monthly turnover and historical volatility, respectively, which are statistically significant at the 1% level. As the dependent and independent variables are standardized, the coefficients imply that a one standard deviation increase in MFD is associated with 8%, 11%, and 26% standard deviation increases in standardized unexplained volume, monthly turnover, and the historical volatility, respectively. Thus, the results in Table 11 provide a strong link between MFD and established measures of investor disagreement.

5.6 Model versus information set disagreement

The investor-specific beliefs in Section 1 are based on differences in the information set $z_{k,i,t}$ and differences in the expectation formation function g_k. In this subsection, we tease out the main driver of overall disagreement with respect to the two aforementioned components. We rewrite Equations (1) and (2) to make each investor k dependent on machine learning model $m, m = 1,...,M$, and information set $j, j = 1,...,J$. Hence, we consider in total M different machine learning models and J different information sets and

Equation (1) is rewritten as

$$E_{k(m,j),t}[r_{i,t+1}] = g_{k(m)}(z_{k(j),i,t}), \tag{6}$$

to stress the dependence of investor k on model m and information set j. Subsequently, we compute disagreement with respect to different models as follows:

$$MFD_{i,t}^{\text{Model}} = \frac{1}{J}\sum_{j=1}^{J}\left(\sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(E_{k(m,j),t}[r_{i,t+1}] - \overline{E_{t,j}[r_{i,t+1}]}\right)^2}\right), \tag{7}$$

where $\overline{E_{t,j}[r_{i,t+1}]} = \frac{1}{M}\sum_{m=1}^{M} E_{k(m,j),t}[r_{i,t+1}]$ is the average forecast across all investors with information set j. Precisely, to calculate model disagreement, we first compute disagreement for each information set j across different models m. Subsequently, we average over all information sets J. Similarly, we write disagreement with respect to different information sets as:

$$MFD_{i,t}^{\text{Information Set}} = \frac{1}{M}\sum_{m=1}^{M}\left(\sqrt{\frac{1}{J}\sum_{j=1}^{J}\left(E_{k(m,j),t}[r_{i,t+1}] - \overline{E_{t,m}[r_{i,t+1}]}\right)^2}\right), \tag{8}$$

where $\overline{E_{t,m}[r_{i,t+1}]} = \frac{1}{J}\sum_{j=1}^{J} E_{k(m,j),t}[r_{i,t+1}]$ is the average forecast across all investors with model m.

In an empirical investigation, we set J equal to 25 and M equal to four, hence, yielding $M \times J = 100$ different investors. The expectation formation function can be either based on random forests with four different sets of hyperparameters or completely different machine learning models by choosing from Lasso, Ridge, random forests, and gradient boosted regression trees.

Table 12 reports results for equal-weighted and value-weighted univariate portfolio sorts. Stocks are cross-sectionally sorted into decile portfolios based on the total MFD, model based MFD, or information set MFD. We report the average long-short portfolio return and the Fama and French (2018) six-factor alpha spreads. Panel A shows results for using different random forest hyperparameters, whereas panel B reports results mixing Lasso, Ridge, random forest, and gradient boosted regression trees as models for the expectation formulation function $g_k(\cdot)$. Table 12 yields several insights on the drivers of our belief disagreement. First, the return spreads based on total MFD are always higher in absolute magnitude compared to either model or information-set-based disagreement. This is probably expected given that total disagreement encompasses two sources of investor disagreement. Second, both model and information-set-based disagreement yield also statistically significant and economically meaningful return spreads and alphas. Third, and most interesting, model based disagreement leads generally to larger return spreads and alphas. This is evident in both cases combining completely different machine learning models or varying the hyperparameter set in random forests. Even though both sources of disagreement contribute to the predictive power of investor disagreement proxied by MFD, our findings suggest that disagreement about the expectation model, that is, how to interpret information, might be a more powerful predictor than disagreement about what information to use.

6. Sources of Return Predictability

Having established a robust negative cross-sectional association between MFD and average stock returns, we next investigate the potential economic mechanisms giving rise to this pattern. Motivated by the theoretical literature and the evidence presented in this paper so far, we explore mispricing versus risk in general, and more specifically investigate limits to arbitrage in the form of short sale constraints and information frictions.

6.1 Mispricing versus risk

In our results so far, we have documented significant alphas controlling for established factor models, a first indication that systematic risks (of the form captured by those models) do not explain the MFD pattern in average returns. Nor can other well-known firm-level risk measures (like idiosyncratic volatility or illiquidity) explain the MFD effect.

If the MFD pattern is indeed associated with mispricing, we expect it to be correlated with other known mispricing phenomena in the literature. In this vein, we compare MFD to the mispricing measure (MISP) of Stambaugh, Yu, and Yuan (2015). We report the time-series average of the cross-sectional mispricing score for stocks in MFD-sorted quintile portfolios.[16] We also conduct dependent double sorts based on individual stock's MISP and MFD; that is, stocks are first grouped into three tercile portfolios on ascending sorts of MISP. Subsequently, stocks are grouped into five quintile portfolios on ascending sorts of MFD within each MISP tercile. We then compute the return spreads and alphas with respect to the Fama and French (2018) six-factor model for MFD high-minus-low portfolios within each MISP quintile.

[16] As discussed in Stambaugh, Yu, and Yuan (2015), each month individual stocks are ranked independently based on 11 prominent equity return predictors (net stock issues, composite equity issues, accruals, net operating assets, asset growth, investment-to-assets, distress, O-score, momentum, gross profitability, and return on assets) associated with lower 1-month-ahead stock returns. The mispricing measure (MISP) is defined as the arithmetic average of the ranks of the 11 return predictors, and a higher (lower) MISP indicates overvaluation (undervaluation).

Table 12
Decomposition of total disagreement into model and information set disagreement

Config IDs	Equal-weighted						Value-weighted					
	Total		Model		Information set		Total		Model		Information set	
	H-L	FF6	H-L	FF6	H-L	FF6	H-L	FF6	H-L	FF6	H-L	FF6
A: Different random forest hyperparameters												
4, 5, 9, 13	−1.03***	−0.69***	−0.93***	−0.56***	−0.90***	−0.62***	−0.92***	−0.64***	−0.73***	−0.57***	−0.68***	−0.43**
	(−4.83)	(−4.98)	(−3.87)	(−3.87)	(−4.94)	(−4.77)	(−4.18)	(−3.30)	(−3.66)	(−3.30)	(−3.38)	(−2.43)
4, 6, 9, 13	−1.02***	−0.69***	−0.92***	−0.54***	−0.90***	−0.62***	−0.90***	−0.60***	−0.76***	−0.57***	−0.67***	−0.43**
	(−4.78)	(−4.88)	(−3.82)	(−3.75)	(−4.88)	(−4.66)	(−4.06)	(−3.12)	(−3.79)	(−3.36)	(−3.16)	(−2.40)
5, 10, 12, 15	−1.27***	−0.85***	−1.11***	−0.63***	−1.17***	−0.78***	−1.07***	−0.62***	−1.01***	−0.50***	−0.89***	−0.45***
	(−4.96)	(−6.86)	(−3.92)	(−4.62)	(−4.82)	(−6.14)	(−3.95)	(−3.67)	(−3.36)	(−2.74)	(−3.29)	(−2.72)
5, 11, 12, 15	−1.17***	−0.80***	−1.03***	−0.57***	−1.06***	−0.67***	−1.08***	−0.66***	−0.88***	−0.49***	−0.85***	−0.44***
	(−4.58)	(−6.24)	(−3.63)	(−4.29)	(−4.40)	(−4.94)	(−4.01)	(−3.75)	(−3.00)	(−2.67)	(−3.15)	(−2.63)
5, 9, 12, 14	−1.31***	−0.96***	−0.99***	−0.64***	−1.13***	−0.75***	−1.27***	−0.90***	−0.86***	−0.49***	−0.85***	−0.44**
	(−5.59)	(−8.74)	(−3.66)	(−5.32)	(−5.20)	(−6.12)	(−5.17)	(−5.20)	(−3.20)	(−2.64)	(−3.54)	(−2.57)
6, 7, 10, 15	−1.11***	−0.79***	−0.88***	−0.51***	−1.06***	−0.66***	−1.04***	−0.67***	−0.83***	−0.51**	−0.87***	−0.43***
	(−4.16)	(−6.29)	(−3.08)	(−3.67)	(−4.22)	(−4.71)	(−3.61)	(−3.98)	(−2.77)	(−2.37)	(−3.04)	(−2.59)
B: Different machine learning models												
1, 11, 17, 18	−1.02***	−0.72***	−0.95***	−0.67***	−0.92***	−0.62***	−0.69***	−0.48***	−0.62***	−0.45***	−0.75***	−0.39**
	(−4.45)	(−5.24)	(−4.40)	(−4.85)	(−4.14)	(−5.55)	(−3.05)	(−3.04)	(−2.90)	(−2.86)	(−2.92)	(−2.33)
1, 3, 17, 18	−1.02***	−0.72***	−0.96***	−0.69***	−0.91***	−0.62***	−0.71***	−0.50***	−0.61***	−0.43***	−0.74***	−0.36**
	(−4.48)	(−5.34)	(−4.49)	(−5.01)	(−4.08)	(−5.49)	(−3.12)	(−3.19)	(−2.85)	(−2.78)	(−2.74)	(−2.06)
1, 7, 17, 18	−1.10***	−0.79***	−1.03***	−0.74***	−0.92***	−0.62***	−0.85***	−0.62***	−0.68***	−0.48***	−0.75***	−0.36**
	(−4.64)	(−5.99)	(−4.60)	(−5.57)	(−3.86)	(−5.12)	(−3.58)	(−3.89)	(−2.95)	(−3.00)	(−2.73)	(−2.04)
2, 16, 17, 18	−1.05***	−0.76***	−1.00***	−0.72***	−0.90***	−0.55***	−0.67***	−0.47***	−0.55***	−0.38**	−0.76***	−0.34**
	(−4.70)	(−6.33)	(−4.61)	(−6.01)	(−3.76)	(−4.12)	(−3.17)	(−3.13)	(−2.69)	(−2.46)	(−2.88)	(−2.04)
2, 6, 17, 19	−1.21***	−0.87***	−1.15***	−0.81***	−1.09***	−0.77***	−0.79***	−0.50***	−0.67***	−0.43***	−0.77***	−0.37**
	(−5.15)	(−7.28)	(−5.07)	(−6.83)	(−4.56)	(−5.97)	(−3.19)	(−3.03)	(−2.89)	(−2.60)	(−3.00)	(−2.14)
2, 8, 17, 19	−1.19***	−0.88***	−1.13***	−0.82***	−1.10***	−0.76***	−0.83***	−0.54***	−0.72***	−0.49***	−0.89***	−0.41**
	(−4.88)	(−7.36)	(−4.76)	(−6.71)	(−4.10)	(−5.62)	(−3.24)	(−3.45)	(−2.96)	(−3.03)	(−2.80)	(−2.20)

The table reports returns and Fama and French (2018) alphas of long-short portfolios formed on MFD. We consider four different versions of the expectation formulation function $g_k(\cdot)$ and 25 unique information sets in Equation (1) yielding in total 100 investors. We measure disagreement across all information sets and models (Total), across the models for the expectation formulation function (Model) or across the information sets (Information set). The models for the expectation formulation function can be either based on random forests with four different sets of hyperparameters (panel A) or by combining the four different machine learning models (panel B): Lasso, Ridge, random forests, and gradient boosted regression trees. Columns labeled H-L and FF6 show average long-short portfolio returns and their Fama and French (2018) alphas, respectively. t-statistics are reported in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Table 13
Mispricing and MFD

A: Average MISP in MFD decile portfolio

	Low	2	3	4	High	H-L	t-stat
MISP	43.94	46.52	48.83	51.35	55.75	11.81***	18.31

B: Bivariate portfolio sort on MISP

	Low	2	3	4	High	H-L	t-stat	FF6	t-stat
MISP low	1.29	1.31	1.20	1.21	0.91	−0.37***	−2.68	−0.31***	−2.76
MISP 2	1.08	1.12	1.09	1.03	0.63	−0.45***	−2.85	−0.33***	−3.17
MISP high	0.73	0.59	0.27	0.25	−0.35	−1.08***	−4.97	−0.68***	−5.10
MISP H-L	−0.56	−0.73	−0.93	−0.95	−1.26	−0.70***	−4.25	−0.37**	−2.51

Panel A reports the time-series averages of the monthly cross-sectional median of the stock-level mispricing score (MISP) of Stambaugh, Yu, and Yuan (2015) for each of MFD-sorted univariate quintile portfolios. Low (high) MISP indicates a lower (higher) mispricing score. Panel B reports 3x5 dependent bivariate equal-weight portfolio sorts. First, tercile portfolios are formed every month using MISP. Next, quintile portfolios are formed based on MFD within each firm-specific MISP tercile. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from August 1976 to December 2016. *p <.1; **p <.05; ***p <.01.

Table 13, panel A, shows that the high MFD stocks indeed have a higher average mispricing score than the low MFD stocks. Furthermore, as reported in the last column of panel A, Table 13, the 5-1 difference in the average mispricing score is 11.81 and statistically significant at the 1% level with a t-statistic of 18.38. Thus, we conclude that high-MFD stocks are more likely to be overvalued.

Next, we investigate whether the cross-sectional relation between MFD and future returns is stronger for overvalued versus undervalued stocks. Specifically, we calculate the return spreads and Fama and French (2018) six-factor alpha spreads of MFD-sorted portfolios within each MISP tercile. Panel B shows results for the equal-weighted bivariate portfolios of MISP and MFD. The last column in panel B presents the FF6 alpha spreads between the high MFD and low MFD quintile portfolios along with the Newey-West t-statistics. A notable point in Table 13 is that the return and alpha spreads on MFD-sorted portfolios increase monotonically (in absolute magnitude) moving from low-MISP to high-MISP tercile, and the FF6 alpha spread is highest at −0.68% per month with a t-statistic of −5.10 for overvalued stocks, that is, in the high MISP tercile. Moreover, the alpha spread on MFD-sorted portfolios of overvalued stocks is economically and statistically greater than the alpha spread on MFD-sorted portfolios for undervalued stocks.

To further differentiate the negative cross-sectional association between MFD and future stock returns from a risk-based explanation, we study stock price reactions around earnings announcements. If the return predictability were explained by underlying risk, we would expect the returns to be evenly affected in subsequent periods. On the contrary, if the effect is consistent with mispricing, then the returns must be disproportionally affected around earnings announcements, that is, the return prediction around earnings announcements should be stronger than that around nonearnings announcement periods if investors are surprised by the good or bad news during that period and revise their expectations. Our approach is widely used in the literature (see, e.g., Bernard and Thomas 1989; Porta et al. 1997; Engelberg, McLean, and Pontiff 2018). We follow Engelberg, McLean, and Pontiff (2018) and conduct a panel regression analysis of daily stock returns (Ret_t^d) on the previous month MFD, an earnings announcement window dummy (EDAY), and the interaction term between the two variables. We also include a set of control variables, consisting of the lagged values for each of the past 10 days for stock returns, stock returns squared, and trading volume. We also control for day fixed effects and cluster the standard errors by day.

The date of the earnings announcement is defined as in Engelberg, McLean, and Pontiff (2018). Specifically, we compute the firm's trading volume scaled by market trading volume for the day before, the day of, and the day after the reported earnings announcement date, which is obtained from Compustat quarterly database. We then define the day with the highest scaled trading volume as the day of the earnings announcement. We select 1-day or 3-day earnings announcement windows centered on the earnings announcement date in our analysis. Panel A of Table 14 reports the regression results for the 1-day window, whereas panel B presents the results for the 3-day window. In all cases and in line with the findings of Engelberg, McLean, and Pontiff (2018), the coefficients on the EDAY are positive and significant. Additionally, the coefficients on MFD are negative and highly statistically significant, corroborating the previously documented negative cross-sectional relation between MFD and future stock returns. More importantly, and consistent with the mispricing explanation, the coefficient for the interaction term between MFD and EDAY is negative and statistically significant, meaning that the negative cross-sectional relation is stronger on earnings announcement days. The coefficient is also economically significant. In column 2 of panel A, the coefficient on MFD is −0.32 (t-stat. = −6.84), while the coefficient of MFD×EDAY interaction term is −0.50 (t-stat. = −3.38), indicating that the return spread for the hedged MFD strategy is 156% higher during an earnings announcement window than on nonannouncement days. Analogously, based on column 4, the MFD premium is 113% higher during a 3-day earnings announcement window than on nonannouncement days. Thus, the evidence supports our mispricing argument that as investors appear to be surprised by the content of new information and subsequently update their beliefs, leading to an elevated MFD-return spread on earnings announcement days.

Table 14
Earnings announcement returns prediction

	A: One-day window		B: Three-day window	
Dep. variable	Ret_t^d	Ret_t^d	Ret_t^d	Ret_t^d
MFD	−0.26***	−0.32***	−0.26***	−0.31***
	(−6.30)	(−6.84)	(−6.15)	(−6.67)
MFD×EDAY	−0.50***	−0.50***	−0.36***	−0.35***
	(−3.39)	(−3.38)	(−5.12)	(−5.06)
EDAY	0.25***	0.26***	0.15***	0.15***
	(9.27)	(9.42)	(11.57)	(11.74)
Lagged controls?	No	Yes	No	Yes
Day fixed effects?	Yes	Yes	Yes	Yes

The table reports results from the panel regressions of daily returns (Ret_t^d) on the previous month's MFD, an earnings announcement window dummy variable (EDAY), an interaction between MFD and EDAY, day fixed effects, and other lagged control variables (coefficients unreported). Ret_t^d, the dependent variable, is multiplied by 100. An earnings announcement window is defined analogously to Engelberg, McLean, and Pontiff (2018) as the 1-day or 3-day window centered on an earnings release, that is, days $t-1$, t, and $t+1$. EDAY is a dummy variable equaling one if the daily observation is during an announcement window, and zero otherwise. Following Engelberg, McLean, and Pontiff (2018), we obtain earnings announcement dates from the Compustat quarterly database and examine the firm's trading volume scaled by market trading volume for the day before, the day of, and the day after the reported earnings announcement date. An earnings announcement day is defined as the day with the highest scaled trading volume. MFD is by construction at the monthly frequency and its previous month value is merged to daily stock returns Ret_t^d. Control variables include lagged values for each of the past 10 days for stock returns, squared stock returns, and trading volume. Standard errors are clustered by day. t-statistics are in parentheses. The sample period is from August 1976 to December 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

6.2 Short-selling costs

In light of the preceding evidence of MFD's association with mispricing, and that this mispricing is most prominent among high disagreement stocks, we investigate the Miller (1977) hypothesis that disagreement combined with short-sale constraints produces overpricing of high MFD stocks. We use two data sets that measure short sale frictions: the indicative borrowing fee provided by IHS Markit, and institutional ownership.

The indicative borrowing fee is calculated from proprietary data by IHS Markit. It is an estimate of the current costs for a hedge fund to borrow shares. Hence, it is regarded as a good proxy for short-sale constraints. Besides, borrowing costs between share lenders and prime brokers, its computation uses also rates from hedge funds to produce an indication of the current market rate. Panel A of Table 15 presents the time-series averages of cross-sectional medians for the indicative fee for equity quintiles formed via a univariate MFD sort. Equities with higher MFD have higher indicative fees (or more binding short-sale constraints), and the difference between the high and the low MFD quintiles is highly significant.

Next, we analyze the strength of the MFD return spread within indicative fee terciles. Panel B of Table 15 shows that the abnormal return (six-factor alpha) to the zero-cost portfolio that buys stocks with the highest MFD and sells stocks with the lowest MFD increases in magnitude from low indicative fee to high indicative fee. For stocks within the lowest

Table 15
Short-sale contraints and MFD

A: Average BORROWFEE in MFD decile portfolio

	Low	2	3	4	High	H-L	t-stat
BORROWFEE	0.67	0.68	0.88	1.27	3.82	3.15***	8.90

B: Bivariate portfolio sort on BORROWFEE

	Low	2	3	4	High	H-L	t-stat	FF6	t-stat
BORROWFEE low	0.92	1.00	1.01	0.88	0.81	−0.11	−0.54	−0.12	−0.85
BORROWFEE 2	1.00	0.92	0.81	0.83	0.18	−0.82***	−2.63	−0.74**	−2.46
BORROWFEE high	0.66	0.38	−0.47	−0.93	−1.96	−2.62***	−5.78	−2.36***	−6.67
BORROWFEE H-L	−0.26	−0.62	−1.48	−1.81	−2.77	−2.51***	−6.39	−2.24***	−6.06

C: Average INST in MFD decile portfolio

	Low	2	3	4	High	H-L	t-stat
INST	0.52	0.52	0.51	0.49	0.42	−0.10***	−9.59

D: Bivariate portfolio sort on INST

	Low	2	3	4	High	H-L	t-stat	FF6	t-stat
INST low	1.16	1.00	0.79	0.43	−0.32	−1.48***	−5.95	−1.10***	−6.89
INST 2	1.11	1.03	0.98	0.78	0.32	−0.79***	−3.57	−0.43***	−3.48
INST high	1.09	0.99	0.93	0.74	0.58	−0.51***	−2.77	−0.31**	−2.49
INST H-L	−0.08	−0.00	0.14	0.32	0.90	0.97***	5.77	0.79***	4.53

Panel A reports the time-series averages of the monthly cross-sectional median of the stock-level indicative borrowing fee (BORROWFEE) taken from IHS Markit for each of MFD-sorted univariate quintile portfolios. Low (high) BORROWFEE indicates a lower (higher) indicate borrowing fee. Panel B reports 3x5 dependent bivariate equal-weight portfolio sorts. First, tercile portfolios are formed every month using BORROWFEE. Next, quintile portfolios are formed based on MFD within each firm-specific BORROWFEE tercile. Panel C reports the time-series averages of the monthly cross-sectional median of the stock-level institutional ownership (INST) for each of MFD-sorted univariate quintile portfolios. Low (high) INST indicates a lower (higher) instutional ownership. Panel D reports 3x5 dependent bivariate equal-weight portfolio sorts. First, tercile portfolios are formed every month using INST. Next, quintile portfolios are formed based on MFD within each firm-specific INST tercile. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from January 2004 to April 2022. * $p < .1$; ** $p < .05$; *** $p < .01$.

tercile (BORROWFEE Low), the FF6 alpha to the zero-cost portfolio is −0.12%, which is not statistically different from zero, while the MFD alpha amounts to −2.36% per month with t-statistic of −6.67 if the indicative fee is highest (BORROWFEE High). The difference in MFD alpha spreads across indicative fee quintiles is economically and statistically significant; −2.24% per month (t-stat.= −6.06). These results indicate that the MFD premium is stronger among stocks with more severe short sale costs as measured by the indicative fee.

Panels C and D in Table 15 repeat the above analysis with an alternative measure of short sale constraints: institutional ownership (Nagel 2005).[17] In panel C of Table 15, we present the time-series averages of cross-sectional means for percentage institutional ownership (INST) for equity quintiles formed via a univariate sort based on MFD. The percentage institutional ownership is equal to 52% for quintile 1. In contrast, for quintile 5, which includes the equities with the highest MFD, the percentage institutional ownership drops to 42%. The difference in institutional holdings between the extreme MFD quintiles is highly significant with a t-statistic of 9.59.

Panel D of Table 15 depicts the strength of the disagreement premium across institutional ownership portfolios using a dependent double sort analysis similar to the above analysis on the indicative borrowing fee. The magnitude of the abnormal return (FF6 alpha) to the zero-cost portfolio that buys stocks with the highest MFD and sells stocks with the lowest MFD increases monotonically in absolute value as one moves toward the stocks for which the level of institutional holdings is lowest (INST low). For stocks in which institutional investors are most active (INST High), the FF6 alpha to the zero-cost portfolio is negative at −0.31% per month (t-stat.=−2.49), whereas the corresponding alpha spread on MFD-sorted portfolios is much higher at −1.10% (t-stat.=−6.89) for stocks in which retail investors are most active (INST Low). The diff-in-diff analysis of the FF6 alpha spreads of the stocks with high versus low institutional holdings also generates an economically and statistically significant difference. Specifically, the difference between the six-factor alphas of the zero-cost MFD-sorted portfolios among the extreme institutional ownership quintiles (INST high – INST low) is 0.79% with a t-statistic of 4.53.

These results once again confirm the Miller (1977) hypothesis that investor disagreement (proxied by MFD) combined with short-sale constraints produces higher degree of overpricing of high-MFD stocks with

[17] Institutional holdings data are obtained from Thompson-Reuters' Institutional Holdings (13F) database. To measure a stock's institutional holdings (INST), we define month-t INST to be the fraction of total shares outstanding that are owned by institutional investors as of the end of the last fiscal quarter during or before month t. Values of INST are available for the period from January 1980 to December 2022.

high short-sale costs, leading to stronger return spreads on MFD-sorted portfolios in the presence of more binding short-sale constraints.

6.3 Limits to arbitrage

In this section, we further explore the role of limits-to-arbitrage. If the predictive power of MFD is driven by mispricing to some extent, then we should expect the return predictability to be more pronounced for stocks with high arbitrage costs. In our next test, we use three proxies of limits-to-arbitrage that are prevalent in the literature.

The prior literature singles out idiosyncratic risk as the primary arbitrage cost (e.g., Pontiff 2006). We rely on Ali, Hwang, and Trombley (2003) and measure the monthly IVOL as the standard deviation of the daily residuals estimated from the regression of the daily excess stock returns on the daily market return over the previous year. Moreover, following Amihud (2002), we use the monthly illiquidity measure as our second proxy, computed as the absolute daily return divided by the daily dollar trading volume, averaged over the last 126 trading days. Finally, we rely on the market capitalization (size) as our third proxy, which is another widely used measure to capture costly arbitrage (e.g., Cohen and Lou 2012; Lee et al. 2019). Instead of using a single proxy for limits-to-arbitrage, we follow Atilgan et al. (2020) and construct a composite index out of the three aforementioned proxies. The arbitrage cost index is created by arranging stocks in ascending order according to their idiosyncratic volatility and their illiquidity. Likewise, stocks are arranged in descending order based on their size. Each stock is assigned a score corresponding to its position in the decile rank for each variable. Finally, the stock-level arbitrage cost index is computed as the sum of these three scores, ranging from 3 to 30. A higher index value indicates more stringent limits-to-arbitrage.

We test the limits-to-arbitrage hypothesis using dependent bivariate portfolios. Specifically, we first sort stocks into tercile portfolios every month based on the arbitrage cost index. Then, we divide each arbitrage cost tercile into quintiles based on MFD. Consistent with the limits-to-arbitrage hypothesis, Table 16 shows that the return and alpha spreads on MFD-sorted portfolios are negative and larger in absolute magnitude, and statistically more significant for stocks with high arbitrage costs, compared to the return and alpha spreads on MFD-sorted portfolios for stocks with low arbitrage costs. The difference of the return and alpha spreads of the stocks with high versus low arbitrage costs also generates a highly significant difference in the MFD premium; the difference in alpha spreads is -0.64% with a t-statistic of -3.30. Thus, we conclude that the slow diffusion of information into stock prices due to limits-to-arbitrage provides a complementary explanation to the predictive power of MFD.

Table 16
Limits-to-arbitrage and MFD

A: Average ARB in MFD decile portfolio

	Low	2	3	4	High	H-L	*t*-stat
ARB	13.40	14.57	15.76	16.82	18.53	5.13***	8.67

B: Bivariate portfolio sort on ARB

	Low	2	3	4	High	H-L	*t*-stat	FF6	*t*-stat
ARB low	1.02	0.92	0.86	0.82	0.60	−0.42***	−4.12	−0.38***	−4.53
ARB 2	1.13	1.07	1.00	0.87	0.41	−0.72***	−4.00	−0.48***	−4.32
ARB high	1.07	0.74	0.50	0.38	−0.44	−1.51***	−6.50	−1.02***	−5.70
ARB H-L	0.05	−0.18	−0.36	−0.44	−1.04	−1.09***	−5.11	−0.64***	−3.30

Panel A reports the time-series averages of the monthly cross-sectional median of a limits-to-arbitrage score (ARB) for MFD-sorted univariate quintile portfolios. Low (high) ARB indicates a lower (higher) average arbitrage cost index. Panel B reports 3x5 dependent bivariate equal-weight portfolio sorts. First, tercile portfolios are formed every month using ARB. Next, quintile portfolios are formed based on MFD within each firm-specific ARB tercile. The arbitrage cost index on the stock-level is constructed using firm size, firm age, idiosyncratic volatility and illiquidity of the stock. To construct it, we sort stocks in increasing order according to their idiosyncratic volatility and illiquidity. Similarly, we sort stocks into decreasing order of firm age and size. Each stock is given the corresponding score of its decile rank for each variable. Finally, the arbitrage cost index on the stock-level is the sum of the four scores such that it ranges from 4 to 40. Newey and West (1987) adjusted t-statistics are reported in parentheses. The sample period is from August 1976 to March 2022. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

7. Conclusion

This paper introduces a statistical model of investor beliefs from which we build a novel measure of investor belief disagreement. In particular, we simulate differences in beliefs across investors by endowing them with different machine learning models for forecasting returns from the same set of inputs. Thus, differences in beliefs across investors emerge from differences in the way they perceive and use data. Investor disagreement is measured as the standard deviation of future return forecasts across investors.

We find a significantly negative and highly robust cross-sectional relation between this newly proposed measure, MFD, and future stock returns. In particular, the value-weighted arbitrage portfolio that takes a short position in the 10th percentile of stocks with the highest MFD and takes a long position in the 10th percentile of stocks with the lowest MFD yields a monthly average return of 1.09%. We also examine the long-term predictive power of MFD and find that the negative relation between MFD and future equity returns persists up to 5 months for the value-weighted portfolios. Finally, we find corroborative evidence for the significance of MFD from bivariate portfolio sorts and multivariate Fama-MacBeth regressions when we control for a large number of firm characteristics and risk factors. As a robustness check, we construct alternative measures of investor disagreement based on the realized next-quarter earnings, instead of next-month excess returns. MFD's predictive power on future stock returns remains significant.

We investigate the source of the MFD spread portfolio's alpha. We conduct comprehensive analyses to differentiate the risk versus mispricing explanations, and present evidence that the alpha for high-MFD stocks is driven primarily by mispricing. To better understand the economic mechanisms behind MFD-based return predictability, we test whether the predictive power of MFD is explained by short-sale constraints and/or other limits to arbitrage. We show that the disagreement premium is significantly stronger for stocks with higher short-sale constraints. Relatedly, the negative relation between MFD and future returns is most pronounced for stocks with high arbitrage costs and high retail ownership. Therefore, our findings support the mispricing explanation of the disagreement premium, consistent with Miller (1977).

Code and Data Availability: The replication code and data are available in the Harvard Dataverse at https://doi.org/10.7910/DVN/BO8SQ8.

Appendix A Control Variables

Table A1
Control variable definitions

Variable	Definition	Source
SIZE	The firm's market capitalization, computed as the market value of outstanding equity at the end of month $t-1$	Fama and French 2008
BM	The firm's book value of equity divided by its market capitalization	Fama and French 2008
STR	The stock's 1-month-lagged return	
MOM	The cumulative return of the stock from month $t-12$ to month $t-1$, omitting the most recent month	Jegadeesh and Titman 1993
OP	The ratio of the firm's operating profits to book equity	Fama and French 2015
AG	The percent growth rate of total assets between 2 consecutive fiscal years	Cooper, Gulen, and Schill 2008
TURN	The turnover of shares during the previous 126 trading days	Datar, Naik, and Radcliffe 1998
ILLIQ	The absolute daily return divided by the daily dollar trading volume, averaged over the last 126 trading days	Amihud 2002
IVOL	The standard deviation of daily residuals from a regression of excess stock returns on the market return over the previous year	Ali, Hwang, and Trombley 2003
SUE	Standardized unexpected earnings	Foster, Olsen, and Shevlin 1984
MAX	The average of the five highest daily returns in month $t-1$	Bali, Cakici, and Whitelaw 2011

The table lists the main control variables used throughout the paper. All data are taken from the data set provided by Jensen, Kelly, and Pedersen (2023).

Appendix B MFD Construction

Table B1
Hyperparameter configuration for MFD construction

Hyperparameter	Value
Investor simulation	
Number of investors (K)	100
Information set size (d_k)	76 (50% of total characteristics)
Random forest model	
Number of trees	2,000
Maximum tree depth	3
Feature subsample fraction	0.05
Observation subsample fraction	0.05
Estimation procedure	
Estimation window	10-year rolling window
Model refitting frequency	Every 12 months [see, e.g., Bali et al. 2023; Gu, Kelly, and Xiu 2020]
In-sample return winsorization	1% at each tail

The table describes the baseline hyperparameter configuration used to generate the primary MFD measure. Winsorization is only applied in-sample, not out-of-sample.

Appendix C Different MFD parameterizations

Table C1
Different hyperparameter parameterizations for MFD

Config ID	Model	Max. depth	Nbr. trees	Subsample	Max. Features	Min. sample split	Feature sample					Min. child weight	lambda
eta	alpha												
							by node	by level	by tree				
1	lasso			1.0									0.0001
2	lasso			1.0									0.01
3	rf	4.0	1,000.0	1.0	0.1	10.0							
4	rf	2.0	2,500.0	1.0	0.15	10.0							
5	rf	3.0	2,000.0	1.0	0.1	5.0							
6	rf	3.0	2,000.0	1.0	0.1	10.0							
7	rf	5.0	2,000.0	1.0	0.1	10.0							
8	rf	5.0	1,000.0	1.0	0.1	10.0							
9	rf	3.0	1,000.0	1.0	0.05	5.0							
10	rf	4.0	2,000.0	1.0	0.05	5.0							
11	rf	4.0	2,000.0	1.0	0.1	5.0							
12	rf	4.0	500.0	1.0	0.05	5.0							
13	rf	3.0	1,000.0	1.0	0.05	10.0							
14	rf	3.0	500.0	1.0	0.05	25.0							
15	rf	5.0	500.0	1.0	0.05	10.0							
16	rf	4.0	1,000.0	1.0	0.15	10.0							
17	ridge			1.0									0.01
18	xgbm	2.0	500.0	0.8			0.4	0.1	0.6	5.0	0.4	0.01	0.06
19	xgbm	3.0	250.0	0.5			0.3	0.1	0.6	10.0	0.25	0.03	0.05

The table reports different parameterizations to construct machine forecast disagreement, MFD, used in Table 12.

References

Agarwal, V., Y. E. Arisoy, and N. Y. Naik. 2017. Volatility of aggregate volatility and hedge fund returns. *Journal of Financial Economics* 125:491–510.

Ali, A., L.-S. Hwang, and M. A. Trombley. 2003. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics* 69:355–73.

Allen, F., S. Morris, and H. S. Shin. 2006. Beauty contests and iterated expectations in asset markets. *Review of Financial Studies* 19:719–52.

Allen, L., T. G. Bali, and Y. Tang. 2012. Does systemic risk in the financial sector predict future economic downturns? *Review of Financial Studies* 25:3000–3036.

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

Anderson, E. W., E. Ghysels, and J. L. Juergens. 2005. Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies* 18:875–924.

Andries, M., M. Bianchi, K. K. Huynh, and S. Pouget. 2025. Return predictability, expectations, and investment: Experimental evidence. *The Review of Financial Studies* 38:1687–729.

Atilgan, Y., T. G. Bali, K. O. Demirtas, and A. D. Gunaydin. 2020. Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns. *Journal of Financial Economics* 135:725–53.

Atmaz, A., and S. Basak. 2018. Belief dispersion in the stock market. *Journal of Finance* 73:1225–79.

Avramov, D., S. Cheng, L. Metzker, and S. Voigt. 2023. Integrating factor models. *Journal of Finance* 78:1593–646.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov. 2009. Dispersion in analysts' earnings forecasts and credit rating. *Journal of Financial Economics* 91:83–101.

Baker, M., and J. Wurgler. 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21:129–51.

Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert. 2023. Option return predictability with machine learning and big data. *Review of Financial Studies* 36:3548–602.

Bali, T. G., N. Cakici, and R. F. Whitelaw. 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99:427–46.

Banerjee, S. 2011. Learning from prices and the dispersion in beliefs. *Review of Financial Studies* 24:3025–68.

Banerjee, S., and I. Kremer. 2010. Disagreement and learning: Dynamic patterns of trade. *Journal of Finance* 65:1269–302.

Barber, B. M., and T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21:785–818.

Barberis, N. 2018. Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 1, 79–175. Elsevier.

Berkman, H., V. Dimitrov, P. C. Jain, P. D. Koch, and S. Tice. 2009. Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements. *Journal of Financial Economics* 92:376–99.

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

Bessembinder, H., K. Chan, and P. J. Seguin. 1996. An empirical examination of information, differences of opinion, and trading activity. *Journal of Financial Economics* 40:105–34.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu. 2006. Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis* 41:455–87.

Boyer, B., T. Mitton, and K. Vorkink. 2010. Expected idiosyncratic skewness. *Review of Financial Studies* 23:169–202.

Carlin, B. I., F. A. Longstaff, and K. Matoba. 2014. Disagreement and asset prices. *Journal of Financial Economics* 114:226–38.

Chan, L. K., J. Karceski, and J. Lakonishok. 2007. Analysts' conflicts of interest and biases in earnings forecasts. *Journal of Financial and Quantitative Analysis* 42:893–913.

Chen, J., H. Hong, and J. C. Stein. 2002. Breadth of ownership and stock returns. *Journal of Financial Economics* 66:171–205.

Chen, L., M. Pelger, and J. Zhu. 2024. Deep learning in asset pricing. *Management Science* 70:714–50.

Cohen, L., and D. Lou. 2012. Complicated firms. *Journal of Financial Economics* 104:383–400.

Cong, L. W., G. Feng, J. He, and X. He. 2025. Growing the efficient frontier on panel trees. *Journal of Financial Economics* 167:104024–.

Cookson, J. A., and M. Niessner. 2020. Why don't we agree? evidence from a social network of investors. *Journal of Finance* 75:173–228.

———. 2023. Investor disagreement: Daily measures from social media. Working paper.

Cooper, M. J., H. Gulen, and M. J. Schill. 2008. Asset growth and the cross-section of stock returns. *Journal of Finance* 63:1609–51.

Couts, S. J., A. S. Gonçalves, and J. Loudis. 2024. The subjective risk and return expectations of institutional investors. Fisher College of Business Working Paper.

Da, Z., X. Huang, and L. J. Jin. 2021. Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics* 140:175–96.

Dahlquist, M., and M. Ibert. 2024. Equity return expectations and portfolios: Evidence from large asset managers. *Review of Financial Studies* 37:1887–928.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52:1035–58.

Daniel, K., D. Hirshleifer, and L. Sun. 2020. Short-and long-horizon behavioral factors. *Review of Financial Studies* 33:1673–736.

Datar, V. T., N. Y. Naik, and R. Radcliffe. 1998. Liquidity and stock returns: An alternative test. *Journal of Financial markets* 1:203–19.

Diether, K. B., C. J. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *Journal of Finance* 57:2113–41.

Doukas, J. A., C. F. Kim, and C. Pantzalis. 2006. Divergence of opinion and equity returns. *Journal of Financial and Quantitative Analysis* 41:573–606.

Dugar, A., and S. Nathan. 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. *Contemporary Accounting Research* 12:131–60.

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

Fama, E. F., and K. R. French. 2008. Dissecting anomalies. *Journal of Finance* 63:1653–78.

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Fama, E. F., and K. R. French. 2018. Choosing factors. *Journal of Financial Economics* 128:234–52.

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

Freyberger, J., A. Neuhierl, and M. Weber. 2020. Dissecting characteristics nonparametrically. *Review of Financial Studies* 33:2326–77.

Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 1189–232.

Garfinkel, J. A. 2009. Measuring investors' opinion divergence. *Journal of Accounting Research* 47:1317–48.

Ge, L., T.-C. Lin, and N. D. Pearson. 2016. Why does the option to stock volume ratio predict stock returns? *Journal of Financial Economics* 120:601–22.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2021. Five facts about beliefs and portfolios. *American Economic Review* 111:1481–522.

Giglio, S., and D. Xiu. 2021. Asset pricing with omitted factors. *Journal of Political Economy* 129:1947–90.

Goetzmann, W. N., and M. Massa. 2005. Dispersion of opinion and stock returns. *Journal of Financial Markets* 8:324–49.

Golez, B., and R. Goyenko. 2022. Disagreement in the equity options market and stock returns. *Review of Financial Studies* 35:1443–79.

Goulding, C. L., C. R. Harvey, and H. Kurtović. 2025. Disagreement of disagreement. Working paper.

Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *Review of Financial Studies* 33:2223–73.

Harris, M., and A. Raviv. 1993. Differences of opinion make a horse race. *Review of Financial Studies* 6:473–506.

Harrison, J. M., and D. M. Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92:323–36.

Hong, H., and J. C. Stein. 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54:2143–84.

———. 2003. Differences of opinion, short-sales constraints, and market crashes. *Review of Financial Studies* 16:487–525.

Hou, K., and R. K. Loh. 2016. Have we solved the idiosyncratic volatility puzzle? *Journal of Financial Economics* 121:167–94.

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

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2015. Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies* 28:791–837.

Jegadeesh, N. 1990. Evidence of predictable behavior of security returns. *Journal of Finance* 45:881–98.

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

Jensen, T. I., B. Kelly, and L. H. Pedersen. 2023. Is there a replication crisis in finance? *Journal of Finance* 78:2465–518.

Jiang, H., and Z. Sun. 2014. Dispersion in beliefs among active mutual funds and the cross-section of stock returns. *Journal of Financial Economics* 114:341–65.

Johnson, T. C. 2004. Forecast dispersion and the cross section of expected returns. *Journal of Finance* 59:1957–78.

Jurado, K., S. C. Ludvigson, and S. Ng. 2015. Measuring uncertainty. *American Economic Review* 105:1177–216.

Kandel, E., and N. D. Pearson. 1995. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103:831–72.

Kelly, B. T., S. Pruitt, and Y. Su. 2019. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134:501–24.

Kumar, A. 2009. Who gambles in the stock market? *Journal of Finance* 64:1889–933.

Laudenbach, C., A. Weber, R. Weber, and J. Wohlfart. 2024. Beliefs about the stock market and investment choices: Evidence from a survey and a field experiment. *Review of Financial Studies hhae063*.

Lee, C. M. C., S. T. Sun, R. Wang, and R. Zhang. 2019. Technological links and predictable returns. *Journal of Financial Economics* 132:76–96.

Li, B., and A. G. Rossi. 2020. Selecting mutual funds from the stocks they hold: A machine learning approach. Working paper.

Loughran, T., and B. McDonald. 2014. Measuring readability in financial disclosures. *Journal of Finance* 69:1643–71.

Malkiel, B. G., and J. G. Cragg. 1970. Expectations and the structure of share prices. *American Economic Review* 60:601–17.

Michaely, R., and K. L. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12:653–86.

Miller, E. M. 1977. Risk, uncertainty and divergence of opinion. *Journal of Finance* 32:1151–68.

Moskowitz, T. J., and M. Grinblatt. 1999. Do industries explain momentum? *Journal of Finance* 54:1249–90.

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

Nagel, S., and Z. Xu. 2023. Dynamics of subjective risk premia. *Journal of Financial Economics* 150:103713–.

Neuhierl, A., X. Tang, R. T. Varneskov, and G. Zhou. 2025. Do option characteristics predict the underlying stock returns in the cross-section? *Management Science*. Advance Access published August 6, 2025, 10.1287/mnsc.2024.04720.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

O'Doherty, M. S., N. E. Savin, and A. Tiwari. 2017. Hedge fund replication: A model combination approach. *Review of Finance* 21:1767–804.

Pontiff, J. 2006. Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics* 42:35–52.

Porta, L., J. L. Rafael, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52:859–74.

Qu, S., L. Starks, and H. Yan. 2003. Risk, dispersion of analyst forecasts and stock returns. Working Paper, University of Texas at Austin.

Scheinkman, J. A., and W. Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111:1183–220.

Sikorskaya, T. 2024. Institutional investors, securities lending and short-selling constraints. Working Paper, University of Chicago Booth School of Business.

Stambaugh, R. F., J. Yu, and Y. Yuan. 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70:1903–48.

Stambaugh, R. F., and Y. Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30:1270–315.

Stein, J. C. 2009. Presidential address: Sophisticated investors and market efficiency. *Journal of Finance* 64:1517–48.

Van Binsbergen, J. H., X. Han, and A. Lopez-Lira. 2023. Man versus machine learning: The term structure of earnings expectations and conditional biases. *Review of Financial Studies* 36:2361–96.

Yu, J. 2011. Disagreement and return predictability of stock portfolios. *Journal of Financial Economics* 99:162–83.

Zhang, X. F. 2006. Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research* 23:565–90.