7-2021

Are disagreements agreeable? Evidence from information aggregation

Dashan HUANG
Singapore Management University, DASHANHUANG@smu.edu.sg

Jiangyuan LI
Shanghai University of Science and Technology

Liyao WANG
Hong Kong Baptist University

Citation
HUANG, Dashan; LI, Jiangyuan; and WANG, Liyao. Are disagreements agreeable? Evidence from information aggregation. (2021). *Journal of Financial Economics*. 141, (1), 83-101.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/6693

Are disagreements agreeable? Evidence from information aggregation[☆]

Dashan Huang[a,*], Jiangyuan Li[b], Liyao Wang[c]

[a] *Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, 178899, Singapore*

[b] *School of Finance, Shanghai University of Finance and Economics, 777 Guoding Rd., Shanghai 200433, China*

[c] *School of Business, Hong Kong Baptist University, 34 Renfrew Road, Kowloon Tong, Kowloon, Hong Kong*

Abstract

Disagreement measures are known to predict cross-sectional stock returns but fail to predict market returns. This paper proposes a partial least squares disagreement index by aggregating information across individual disagreement measures and shows that this index significantly predicts market returns both in- and out-of-sample. Consistent with the theory in Atmaz and Basak (2018), the disagreement index asymmetrically predicts market returns with greater power in high-sentiment periods, is positively associated with investor expectations of market returns, predicts market returns through a cash flow channel, and can explain the positive volume-volatility relationship.

Keywords: Disagreement, Return predictability, PLS, LASSO, Machine learning

JEL classification: G12, G14, G17

☆ We are grateful to G. William Schwert (the editor) and Seth Pruitt (the referee) for very insightful and helpful comments that improved the paper significantly. We thank Adem Atmaz, Suleyman Basak, Bong-Geun Choi, Liyuan Cui, Zhi Da, Stefano Giglio, Shiyang Huang, Tao Li, Weikai Li, Ye Li, Francis Longstaff, Sungjune Pyun, and Guofu Zhou as well as seminar and conference participants at St. Louis University, Washington University in St. Louis, 2018 Asian Bureau of Finance and Economic Research (ABFER) Annual Conference, 2018 AsianFA Annual Meeting, 2018 City University of Hong Kong International Finance Conference on Corporate Finance and Financial Markets, 2018 Research in Behavioral Finance Conference, and 2018 SMU Finance Summer Camp for insightful comments. Dashan Huang acknowledges that this study was partially funded at the Singapore Management University through a research grant (MSS18B004) from the Ministry of Education Academic Research Fund Tier 1.

* Corresponding author.

Email addresses: `dashanhuang@smu.edu.sg` (Dashan Huang), `lijiangyuan@mail.shufe.edu.cn` (Jiangyuan Li), `lywang@hkbu.edu.hk` (Liyao Wang)

1. Introduction

Researchers in economics and finance have long been interested in studying the effects of expectations across investors. Investor disagreement, usually measured by the second moment of investor expectations, plays an important role in explaining stock returns, volatility, and trading volume. Due to its wide impacts, Hong and Stein (2007) conclude that disagreement represents "the best horse" for behavioral finance to obtain as many insights as classical asset pricing theories. However, unlike Baker and Wurgler's (2006) sentiment index that has been widely used to capture the first moment of investor expectations (see, e.g., Yu and Yuan, 2011; Stambaugh, Yu, and Yuan, 2012), investor disagreement has only been approximated through various proxies in the literature.[1] To date, there is a lack of research that examines disagreement measures collectively and it is unclear whether they are able to predict market returns in real time.

This paper examines whether extant disagreement measures can become agreeable. If extant measures capture disagreement, they should display commonality and have a common factor. To aggregate information across 24 individual measures, we propose a disagreement index by using the partial least squares (PLS) method in Kelly and Pruitt (2013, 2015). Empirically, we show that the 24 individual measures do have a common factor and the disagreement index significantly predicts market returns up to 12 months. Over the sample period of 1969:12–2018:12, a one-standard-deviation increase in the disagreement index is associated with a 0.83% decrease in the next one-month market return and a 7.04% decrease in the next 12-month market return, where the latter is comparable to 6.6% in Yu (2011), who measures investor disagreement with analyst forecast dispersion. The in- and out-of-sample R^2s are 2.52% and 1.56% at the one-month horizon and 13.88% and 13.26% at the 12-month horizon. In contrast, there are only four individual disagreement measures that are significant at the one-month horizon and four others significant

[1] Professional forecast dispersions (Anderson, Ghysels, and Juergens, 2009; Li, 2016; Bordalo, Gennaioli, Ma, and Shleifer, 2020), analyst forecast dispersions (Diether, Malloy, and Scherbina, 2002; Hong and Sraer, 2016), household forecast dispersions (Li and Li, 2017), unexplained trading volume (Garfinkel, 2009), and stock idiosyncratic volatility (Boehme, Danielsen, and Sorescu, 2006) are some of the most prominent disagreement measures to date.

at the 12-month horizon for in-sample forecasting, but none of the 24 individual measures exhibits any out-of-sample forecasting power.

PLS is chosen for information aggregation due to its simplicity and efficacy. Initially proposed by Wold (1966) and further developed by Kelly and Pruitt (2013, 2015), it extracts the disagreement index with a three-pass regression filter to reduce common noises in the individual disagreement measures. Theoretically, PLS outperforms principal component analysis (PCA) in extracting factors for prediction if individual predictors contain a common (noise) component that is unrelated to future market returns. The intuition is that, as a supervised learning technique, PLS incorporates the target information—market returns—in the factor extracting procedure and teases out any common component that is uncorrelated with future market returns. Empirically, Kelly and Pruitt (2013), Lyle and Wang (2015), Huang, Jiang, Tu, and Zhou (2015), Giglio, Kelly, and Pruitt (2016), Light, Maslov, and Rytchkov (2017), and Gu, Kelly, and Xiu (2020), among others, show that PLS is effective at extracting factors for predicting stock returns and economic activities in the time series and cross-section.

The forecasting power of the disagreement index is not subsumed by economic predictors and uncertainty measures. It remains significant after controlling for the 14 economic predictors in Welch and Goyal (2008), output gap in Cooper and Priestley (2009), and aggregate short interest in Rapach, Ringgenberg, and Zhou (2016). Also, while the disagreement index represents one type of uncertainty (see, e.g., Anderson, Ghysels, and Juergens, 2009; Atmaz and Basak, 2018), it is distinct from extant uncertainty measures, such as economic uncertainty (Bali, Brown, and Caglayan, 2014), treasury implied volatility (Choi, Mueller, and Vedolin, 2017), financial uncertainty and macro uncertainty (Jurado, Ludvigson, and Ng, 2015), economic policy uncertainty (Baker, Bloom, and Davis, 2016), news implied volatility (Manela and Moreira, 2017), sample variance (Welch and Goyal, 2008), and the Chicago Board Options Exchange (CBOE) volatility index (VIX).

The ability of disagreement to predict market returns is robust to alternative econometric methods. In addition to PLS, we explore six LASSO-related machine learning methods (see, e.g., Rapach, Strauss, and Zhou, 2013; Chinco, Clark-Joseph, and Ye, 2019; Diebold and Shin, 2019; Han, He, Rapach, and Zhou, 2020; Freyberger, Neuhierl, and Weber, 2020; Kozak, Nagel, and Santosh, 2020), and find that all of them generate significant out-of-sample R^2s, although the magnitudes are slightly smaller than that with the PLS disagreement index. For example, using the elastic net, the out-of-sample R^2 is 1.36% at the one-month horizon and 8.43% at the 12-month horizon, respectively; both are significant at the 5% level. These results suggest genuine predictability of extant disagreement measures on market returns.

After providing evidence on the forecasting power of the disagreement index, we show that it is indeed consistent with the theory in Atmaz and Basak (2018). In their equilibrium model with infinite heterogeneous investors, Atmaz and Basak (2018) show that the overall effect of belief heterogeneity depends on two sufficient statistics, average bias and disagreement, which can be intuitively defined as the mean and cross-sectional standard deviation of investor expectation biases. Suppose investors are risk averse and exhibit a wealth effect that endogenously limits their risk taking. Atmaz and Basak (2018) show that, in equilibrium, investor disagreement affects stock returns via two channels. The first channel is a direct effect: disagreement represents uncertainty and investors require a higher expected return to hold a stock when disagreement on the stock increases, suggesting a positive disagreement-return relation. The second channel is an indirect effect: investor disagreement affects stock returns via an amplification effect on the average bias. That is, higher disagreement leads to higher average bias and more overvaluation, thereby implying a negative disagreement-return relation. With these two channels, Atmaz and Basak (2018) reconcile the mixed disagreement-return relation shown in the empirical finance literature (see, e.g., Chen, Hong, and Stein, 2002; Diether, Malloy, and Scherbina, 2002; Yu, 2011; Carlin, Longstaff, and Matoba, 2014). Since investors, regardless of whether they are sophisticated or not, are generally upward biased (see, e.g., Barber and Odean, 2008; Edelen, Ince, and Kadlec, 2016; DeVault, Sias, and Starks, 2019; Engelberg, McLean, and Pontiff, 2020), the second channel is more likely to dominate the first channel, which explains why the disagreement index negatively predicts market returns in this paper.

We now test four implications raised by Atmaz and Basak (2018). The first, and most important, implication is that the forecasting power of disagreement is asymmetric: it is stronger when investors are optimistic or among stocks with optimistic investor expectations, and weaker or insignificant otherwise. The intuition is that when investors are relatively pessimistic, the first and second channels have different forecasting signs and are likely to offset each other, making the disagreement-return relation insignificant. In contrast, when investors are overly optimistic, the second channel dominates the first channel, and, as a consequence, disagreement negatively predicts future stock returns. To test this implication, we perform two tests. First, in time series, we show that the forecasting power of disagreement is concentrated in high-investor sentiment periods and nonexistent in low-sentiment periods. Second, cross-sectionally, we form ten decile portfolios based on firm level investor expectation, which is measured by the analyst long-term growth rate (LTG) forecast (Bordalo, Gennaioli, La Porta, and Shleifer, 2019), and find that the disagreement index displays much stronger power in predicting portfolios with higher LTG forecasts, especially in high-sentiment periods.

The second implication is that disagreement should be linked to investor optimism about market returns and ex post forecast errors. To capture investor expectation of market returns, we consider four measures, including aggregate analysts' return forecast (Engelberg, McLean, and Pontiff, 2020), the University of Michigan Surveys of Consumers Attitudes (Das, Kuhnen, and Nagel, 2019), Graham-Harvey's survey of CFOs, and Shiller's survey of individual investor confidence (Greenwood and Shleifer, 2014). We find that all these measures positively correlate with the disagreement index. For example, a one-standard-deviation increase in disagreement is associated with a 3.26% increase in the analysts' return forecast about the following 12-month market return. Since investor expectations are upward biased, the disagreement index negatively predicts ex post return forecast errors.

The third implication is that the predictive ability of disagreement on market returns is more likely to operate via a cash flow channel in the sense of Campbell (1991). According to Atmaz and Basak (2018), after positive cash flow news, investors whose beliefs are supported by the cash flow news become relatively wealthier, which makes them more optimistic about future cash flows or discount rates or both and, consequently, increases investor disagreement. For this reason, both the cash flow news and discount rate news can have a positive effect on disagreement. Empirically, we find that the cash flow news-based disagreement index displays strong forecasting power, while the discount rate news-based disagreement index does not.

The fourth, and last, implication is that disagreement plays an important role in the positive relation between trading volume and market volatility. In Atmaz and Basak (2018), disagreement is the only driver of trading volume and market volatility. In the absence of disagreement, there is no trade and market volatility is constant. In the presence of disagreement, however, both trading volume and market volatility increase as disagreement increases. Empirically, we find that the disagreement index is positively related to the volume-volatility elasticity. Intuitively, a one-standard-deviation increase in disagreement predicts a 5.22% increase in the volume-volatility correlation in the following month. Overall, our empirical results are consistent with the theoretical implications of Atmaz and Basak (2018).

This paper contributes to the disagreement literature by showing that disagreement predicts market returns in- and out-of-sample. While many papers have explored the relation between disagreement and stock returns at the firm level, studies at the market level are relatively rare. There are two exceptions, Yu (2011) on the stock market and Carlin, Longstaff, and Matoba (2014) on the mortgage market, but they do not investigate the out-of-sample forecasting power and the economic value for a real time investor. Also, Yu (2011) documents a negative forecasting sign, whereas Carlin, Longstaff, and Matoba (2014) find a positive forecasting sign; therefore, they interpret their results with different theories. This paper reconciles the seemingly conflicting results by using the unified theory of Atmaz and Basak (2018).

This paper is also related to the broad literature on return predictability. Since Welch and

Goyal (2008), a large number of variables have been identified that significantly predict market returns in- and out-of-sample. They include the output gap (Cooper and Priestley, 2009), 52-week high and historical high (Li and Yu, 2012), aggregate implied cost of capital (Li, Ng, and Swaminathan, 2013), disaggregate book-to-market ratio (Kelly and Pruitt, 2013), aggregate short interest (Rapach, Ringgenberg, and Zhou, 2016), aggregate liquidity (Chen, Eaton, and Paye, 2018), fourth quarter consumption (Møller and Rangvid, 2015), metal prices (Jacobsen, Marshall, and Visaltanachoti, 2019), dividend-price ratio (Golez and Koudijs, 2018), variance risk premium (Pyun, 2019), gold-to-platinum price ratio (Huang and Kilic, 2019), aggregate skewness (Jondeau, Zhang, and Zhu, 2019), and many others. This paper does not aim to identify a new variable to predict market returns, but proposes to aggregate predictive information from extant individual disagreement measures.

The rest of the paper is organized as follows. Section 2 considers 24 extant disagreement measures and shows that they fail to predict the stock market at the one- to 12-month horizons. Section 3 proposes a PLS disagreement index constructed by aggregating information across individual measures and shows that it significantly predicts market returns in- and out-of-sample. Section 4 shows that the predictability of disagreement on market returns is consistent with the theoretical implications of Atmaz and Basak (2018). Section 5 makes a brief conclusion.

2. Forecasting power of extant disagreement measures

At the one- to 12-month horizons, we show in this section that most of the extant disagreement measures fail to predict market returns in-sample and none of them displays significant out-of-sample forecasting power.

2.1. Extant disagreement measures

We consider 24 disagreement measures, among which 13 are based on professional forecasts on eight macro variables, two on analyst forecasts, six on household forecasts on macroeconomic conditions, and three on market information. While these measures originate in different time periods, dating as early as December 1968, all of them conclude in December 2018.

2.1.1. Thirteen disagreement measures based on professional forecasts

The disagreements between professional forecasts on macro variables are based on the Survey of Professional Forecasters (SPF), the oldest quarterly survey of macroeconomic forecasts in the United States. The survey began in 1968Q4 and is typically released in the mid-to-late second month of each quarter.[2] However, the accurate release dates before 1990Q2 are unavailable; therefore, to be conservative, we assume that all surveys are available in the last month of each quarter in our analysis. Also, because most of our analyses are performed monthly, we convert the quarterly measures into monthly frequencies by assigning the most recent quarterly value to each month. For example, the observation in the first quarter of 2018 is assigned to the months of March, April, and May, respectively.

We consider professional forecasts on eight macro variables, including gross domestic production (GDP), industrial production (IP), consumption (CON), investment (INV), housing starts (HSG), unemployment (UEP), the Consumer Price Index (CPI), and the 3-month T-bill rate (TBL). As the forecasts on GDP, IP, CON, INV, and HSG include both level and growth rate, we therefore have 13 disagreement measures in total. In each quarter, the forecasters predict macro variables for horizons ranging from the current up to four quarters ahead. Following Li (2016) and documents from the SPF, we define disagreement on each macro variable as the difference between the 75th percentile and 25th percentile forecasts for each horizon, taking the average across all horizons as the disagreement measure of that macro variable. In the literature, Anderson, Ghysels, and Juergens (2009) and many others use the SPF in a similar fashion in constructing aggregate uncertainty and disagreement measures, and find significant power for pricing the cross-section of stock returns.

[2]Three exceptions with delayed releases are 1990Q2, 1996Q3, and 2013Q4, respectively.

2.1.2. *Two disagreement measures based on analyst forecasts*

Numerous studies have employed analyst forecast dispersion as the measure of investor disagreement. Following Yu and Yuan (2011) and Hong and Sraer (2016), we adopt the "bottom-up" approach by defining disagreement in month t as:

$$D_t^{\mathrm{Yu}} = \frac{\sum_i \mathrm{MKTCAP}_{i,t} \cdot D_{i,t}}{\sum_i \mathrm{MKTCAP}_{i,t}}, \tag{1}$$

and

$$D_t^{\mathrm{HS}} = \frac{\sum_i \beta_{i,t} \cdot D_{i,t}}{\sum_i \beta_{i,t}}, \tag{2}$$

where $D_{i,t}$ is the analyst forecast dispersion on the earnings per share (EPS) long-term growth rate (LTG) of firm i, and MKTCAP and $\beta_{i,t}$ are firm i's market cap and market beta. We only include common stocks (with CRSP item SHRCD = 10 or 11) listed on the NYSE, NASDAQ, and AMEX. As explained in Yu (2011), the LTG forecast features prominently in valuation models and is less affected by a firm's earnings guidance than the short-term forecast. When constructing D_t^{HS}, we follow Hong and Sraer (2016) and focus on all-but-micro stocks, stocks that are larger than the 20th percentile of the market cap of NYSE stocks. For each firm i in month t, we regress the daily returns of the past one year on contemporaneous and one to five lagged market returns, and use the sum of the slopes as the estimate of $\beta_{i,t}$.

2.1.3. *Six disagreement measures based on household forecasts*

Empirical studies often focus on how the trading of securities is affected by disagreement among institutional investors (see, e.g., Diether, Malloy, and Scherbina, 2002; Chen, Hong, and Stein, 2002; Jiang and Sun, 2014), but seldom explore the disagreement effect of households or retail investors on the stock market. From the Flow of Funds Accounts, households own about 60% of outstanding equities in the United States (about 40% direct holding and an additional

20% indirect holding through mutual funds) and, therefore, their opinions should play a similarly important role as those of institutional investors. Li and Li (2017) show that the effect of household disagreement remains significant after controlling for professional forecast dispersions and even dominates the professional forecast dispersion measures.

We construct household disagreement based on the University of Michigan Surveys of Consumers Attitudes (SCA). The SCA started conducting monthly surveys on a minimum of 500 households in January 1978, with accurate release dates available after January 1991. In each survey, the SCA collects responses to 50 core questions that are generally related to households' opinions about current economic conditions and their expectations about future economic conditions. In this paper, we construct our disagreement measures from six questions. The first question is about households' realized opinions about their current personal financial condition compared with those of the prior year, while the other five are about households' expectations about the following year, including their expected personal financial condition, business condition, unemployment condition, interest rate condition, and house purchase condition.

For each question, the surveyed households' replies are classified into three categories: better (good), same (depends), and worse (bad). In a consistent way, we rename the categories as positive, neutral, and negative, respectively, and define the proportion of each category as P_{positive}, P_{neutral}, and P_{negative}. We follow Li and Li (2017) and define disagreement as the unevenly weighted negative Herfindahl index,

$$D = -\sum w_i P_i^2, \quad i = \text{positive, neutral, negative,} \tag{3}$$

where w_i is the weight of each category as $w_{\text{positive}} = 1$, $w_{\text{neutral}} = 2$, and $w_{\text{negative}} = 1$. We assign a higher weight to the neutral category to avoid the unfavourable feature of the evenly weighted Herfindahl index. For example, if 50% of households indicate the positive response and 50% indicate the negative response, the evenly weighted Herfindahl index would be the same as if the responses were 50% positive and 50% neutral. However, the disagreement in the former situation is obviously more dispersed than in the latter.

2.1.4. *Disagreement based on unexplained stock trading volume*

Ajinkya, Atiase, and Gift (1991) find that high trading volume is associated with an increase in the analyst forecast dispersion, suggesting that trading volume could measure investor disagreement. We follow Garfinkel (2009) and construct a disagreement measure with the standardized unexplained volume. Specifically, we obtain the monthly aggregate trading volume data of the NYSE from Pinnacle and define volume as the residual of applying an AR(4) to the log turnover with the past 120-month observations (Hamilton, 2018). Then, we run the following time series regression with data from the past 120-month period at the end of each month on a rolling basis as

$$\text{Volume}_t = \alpha + \beta_1 R_t^+ + \beta_2 R_t^- + \varepsilon_t, \tag{4}$$

and use the last value of the residuals as the estimate of unexpected volume. In Eq. (4), the plus and minus signs in the superscript indicate that market returns can be either positive or negative, and capture the empirical fact that positive and negative returns generate different levels of trading volume. Thus, investor disagreement can be defined by the standardized unexplained volume:

$$D_t^{\text{SUV}} = \frac{\varepsilon_t}{\sigma_{\varepsilon,t}}, \tag{5}$$

where $\sigma_{\varepsilon,t}$ is the standard deviation of the regression residuals.

2.1.5. *Disagreement based on idiosyncratic volatility*

Inspired by theoretical studies that suggest a positive relation between belief dispersion and volatility, Boehme, Danielsen, and Sorescu (2006) and Berkman, Dimitrov, Jain, Koch, and Tice (2009) propose idiosyncratic volatility as a disagreement measure at the firm level. We extend this measure to the market level. Specifically, following Ang, Hodrick, Xing, and Zhang (2006), we regress daily stock returns on the Fama and French (1993) three factors with a 12-month rolling window and estimate the firm level idiosyncratic volatility at the end of each month. We then define investor disagreement as the value-weighted idiosyncratic volatility.

2.1.6. Disagreement based on option open interest

Disagreement can also be constructed from the option market. Investors who hold call options have a bullish view, whereas investors who hold put options have a bearish view. Following Ge, Lin, and Pearson (2016), we define disagreement as one minus the scaled difference between the OEX call and put open interests:

$$D_t^{\mathrm{OID}} = 1 - \frac{|\mathrm{COI}_t - \mathrm{POI}_t|}{|\mathrm{COI}_t + \mathrm{POI}_t|}, \tag{6}$$

where COI_t (POI_t) is the call (put) option open interest. The scaled call and put option open interest difference $|\mathrm{COI}_t - \mathrm{POI}_t|/|\mathrm{COI}_t + \mathrm{POI}_t|$ ranges from zero to one. The explanation is that when disagreement is low, investors' beliefs polarize into bullish or bearish extremes. The difference between the call and put option open interests diverges and the scaled difference approaches one. As a result, one minus this scaled difference is accordingly low. When disagreement is high, the opinions between optimists and pessimists diverge. The call and put option open interests should be commensurable. The scaled difference between the call and put option open interests approaches zero. Hence, one minus the scaled difference is accordingly large.

2.2. Summary statistics

Table 1 presents summary statistics of the 24 disagreement measures, including the sample period, mean, standard deviation, minimum, maximum, skewness, and kurtosis. It is apparent that the scales across disagreement measures vary dramatically due to the nature of macro variables. For instance, the mean of disagreement on GDP is 61.32 billion, while the mean of disagreement on TBL is only 0.46%. Thus, to make them comparable and to avoid forward-looking bias, we standardize each disagreement measure in month t by its last six-year mean and standard deviation, with a requirement of at least one year of data. For this reason, the analyses in all other tables start in December 1969. To remove possible fundamental information, we measure disagreement as the residuals from the regression of each individual disagreement measure on the six macro variables in Baker and Wurgler (2006), consisting of the growth of industrial production, the growth of durable consumption, the growth of nondurable consumption, the growth of service consumption, the growth of employment, and a dummy variable for NBER-dated recessions (we recursively do so when performing out-of-sample tests).

Table A1 in the Online Appendix presents pairwise correlations between individual disagreement measures. Most of the measures are positively correlated, with the exception of several negative values. For example, professional forecast dispersions are generally positively correlated, and they are also positively correlated with the two analyst forecast dispersion measures. Business condition forecast dispersion is an exception, and it is negatively correlated with other measures in general. Overall, this table indicates that extant measures capture both the common and different aspects of individual disagreement measures across the whole economy, and an individual measure is unlikely to completely capture the aggregate effect of disagreement on the stock market.

2.3. Forecasting market returns with extant disagreement measures

We explore the forecasting power of disagreement on market returns with the following predictive regression,

$$R_{t+1} = \alpha + \beta D_t + \varepsilon_{t+1}, \tag{7}$$

where R_{t+1} is the log excess return of the S&P 500 index in month $t+1$ and D_t is one of the 24 individual disagreement measures.[3] When the forecast horizon is h months, we denote the cumulative market return as $R_{t,t+h} = \sum_{j=1}^{h} R_{t+j}$.

[3]For brevity, returns in this paper always refer to excess returns except in Section 4.3, where we follow Campbell (1991) and decompose the total market returns into three components.

The predictive power is assessed based on the regression slope β or the R^2 statistic. If β is significantly different from zero or if the R^2 is significantly larger than zero, it means that D_t is a predictor of the market returns. The out-of-sample forecast of the next one-month market return is recursively computed as

$$\hat{R}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t D_t, \tag{8}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the ordinary least squares estimates of α and β based on data from the start of the available sample through month t. The in-sample forecast is computed as before, except that $\hat{\alpha}_t$ and $\hat{\beta}_t$ are replaced by those estimated by using the entire sample. For ease of exposition, we always normalize the time series of disagreement in all the in-sample predictive regressions, so that the regression slope measures the change in response to a one-standard-deviation increase in disagreement.

We use the out-of-sample R^2 statistic in Campbell and Thompson (2008) as the out-of-sample performance evaluation criterion and define it as:

$$R^2_{OS} = 1 - \frac{\sum_{t=M+1}^{T} (R_t - \hat{R}_t)^2}{\sum_{t=M+1}^{T} (R_t - \bar{R}_t)^2}, \tag{9}$$

where M is the size of in-sample parameter training and $T - M$ is the number of out-of-sample observations. $\hat{R}_t$ is the market return forecast with Eq. (8), and $\bar{R}_t$ is the historical return mean, both of which are estimated using data up to month $t-1$. If D_t is a valid predictor, its mean squared forecast error (MSFE) is lower than the MSFE with the historical return mean and the R^2_{OS} will be positive. Campbell and Thompson (2008) show that a monthly R^2_{OS} of 0.5% can generate a significant economic value. The null hypothesis of interest is therefore $R^2_{OS} \leq 0$ against the alternative hypothesis that $R^2_{OS} > 0$. We test this hypothesis by using the MSFE-adjusted statistic, as proposed by Clark and West (2007).

The left panel of Table 2 presents the regression slope β, Newey-West t-value, in-sample R^2, and out-of-sample R^2_{OS}. Throughout this paper, the out-of-sample period is from February 1991 to December 2018 because the accurate release dates of household dispersion measures are only available as of January 1991. Twenty out of 24 disagreement measures have a negative forecasting sign, however, only four of them reveal significant in-sample predictive power at the 5% level; they are the housing starts forecast dispersion, CPI forecast dispersion, TBL forecast dispersion, and business condition forecast dispersion. The out-of-sample performance is more dismal, with all R^2_{OS} values being negative. For instance, the TBL forecast dispersion exhibits the highest in-sample R^2 of 1.94%, but generates a -4.21% out-of-sample R^2_{OS}. These results suggest that none of the extant individual disagreement measures can predict market returns in real time at the one-month forecast horizon.

The middle and right panels of Table 2 present results similar to the left panel when the forecast horizon is extended to three or 12 months. The in-sample regression slopes are seldom significant and the R^2_{OS} values are all negative. For in-sample prediction over the 1981:12–2005:12 sample period, Yu (2011) shows that analyst forecast dispersion exhibits insignificant forecasting power at the one-month horizon but significant forecasting power at the 12-month or longer horizons. The right panel suggests that when we extend the sample to the most recent period, analyst forecast dispersion becomes insignificant. Yu (2011) does not show out-of-sample forecasting performance and our results suggest that analyst forecast dispersion cannot generate meaningful real time forecasting value either.

Overall, Table 2 shows that while all of the extant disagreement measures could have cross-sectional forecasting power, they are unable to predict market turns in general, especially for out-of-sample forecasting.

3. PLS disagreement index

In this section, we construct a disagreement index by aggregating information across individual disagreement measures and show that it significantly predicts market returns in- and out-of-sample.

3.1. Methodology

The method we choose for information aggregation is PLS, which consists of three steps. In the first step, we run a time series regression of each individual disagreement measure on the realized subsequent market returns (as a proxy of expected return) with the full sample, denoted as:

$$D_{t-1}^k = \pi_{k,0} + \pi_k R_t + u_{k,t-1}, \quad k = \text{GDP}, \cdots, \text{ OID}, \tag{10}$$

where π_k captures the sensitivity of proxy D_{t-1}^k to the expected market return. In the second step, we run a cross-sectional regression of D_t^k on π_k at the end of each month:

$$D_t^k = a_t + D_t \pi_k + \nu_{k,t}, \tag{11}$$

where the regression slope D_t is the PLS disagreement index in month t. In the third step, to predict R_{t+1}, we run the following predictive regression:

$$R_{t+1} = \alpha + \beta D_t + \varepsilon_{t+1}. \tag{12}$$

The above three steps are for in-sample analysis. For out-of-sample forecasting, the standard approach is to repeat the three steps by truncating the observations that are not known at month $t+1$. Specifically, consider a forecast for return R_{t+1} that is realized in month $t+1$. A properly constructed forecast can only use information known through month t. In the first step, the latest return that can be used on the right-hand side is R_t and the last observation of disagreement on the left-hand side is, therefore, D_{t-1}^k. In the second step, the cross-sectional regressions are run from months 1 through t. In the last step, the latest return on the left-hand side entering the predictive regression is R_t and the forecast for R_{t+1} is $\hat{\alpha}_t + \hat{\beta}_t D_t$, where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the estimates using information up to month t. In summary, for out-of-sample forecasting, all inputs to the forecast are constructed using data that are observed no later than month t.

To iron out extreme outliers, we smooth the disagreement index with its six-month moving average values and plot the time series in Fig. 1. There are two interesting observations. First, the disagreement index is time-varying and does not diminish over time, which is consistent with the finding in Hong and Stein (2007) and Cookson and Niessner (2020) that permanent disagreement can arise even when investors have common priors and observe the same time series of public information, so long as they interpret information differently. Second, the disagreement index value can be large in bad times, such as the recessions of 1981 to 1982 and 2007 to 2008, and also in good times, such as the dot-com boom of the late 1990s. This evidence is consistent with the beta-weighted analyst forecast dispersion in Hong and Sraer (2016).

3.2. Forecasting performance

This section explores the in- and out-of-sample forecasting performance of the disagreement index. For comparison, we consider two alternative disagreement indexes as benchmarks. The first alternative disagreement index is constructed based on PCA, which extracts the first principal component of the 24 individual disagreement measures as the aggregate index. This method has been widely used in finance; Baker and Wurgler (2006), for example, construct an investor sentiment index as the first principal component of six individual sentiment proxies. The second alternative disagreement index is constructed by simply equal-weighting the 24 (standardized) individual disagreement measures. The intuition is that if each individual measure is unbiased, equal-weighting will efficiently reduce the idiosyncratic errors.

Panel A of Table 3 reports the results of predicting market returns with the three disagreement indexes. At the one-month horizon, a one-standard-deviation increase in disagreement leads to a 0.38% decrease in the next one-month market return with the PCA disagreement index (t-value $= -1.96$), a 0.60% decrease with the equal-weight disagreement index (t-value $= -2.87$), and a 0.83% decrease with the PLS disagreement index (t-value $= -3.96$). When turning to out-of-sample forecasting, the R^2_{OS} with the PCA disagreement index is 0.20% and not significant. In contrast, the R^2_{OS} is 0.90% with the equal-weight disagreement index and 1.56% with the PLS disagreement index, which are both significant at the 5% level.

Panels B and C of Table 3 report the results when the forecast horizons are three and 12 months, respectively. In these two cases, all the three disagreement indexes display significant in- and out-of-sample forecasting power. For example, a one-standard-deviation increase in disagreement leads to 2.92%, 4.93%, and 7.04% decreases in the next 12-month market returns with the three disagreement indexes, respectively. The R^2 and R^2_{OS} of the PLS disagreement index, 13.88% and 13.26%, are comparable with the most powerful predictor to date, the aggregate short interest in Rapach, Ringgenberg, and Zhou (2016), whose corresponding values are 12.89% and 13.24%, respectively.

Why does the PLS disagreement index have stronger forecasting power than the two alternative disagreement indexes? The reason is that while the PCA and equal-weight disagreement indexes can efficiently reduce the idiosyncratic measurement and observation errors in the individual disagreement measures, they cannot tease out the common errors that are unrelated to expected market returns. In contrast, as a supervised learning technique, the PLS aggregates information relevant to expected market returns and is supposed to perform the best.

To better understand their differences in forecasting power, Fig. 2 depicts the forecasted three-month market returns based on the PCA, equal-weight, and PLS disagreement indexes for the 1991:02–2018:12 out-of-sample period (the results with other forecasting horizons are similar and omitted for brevity). The PLS disagreement index generates more volatile forecasts than the other two and naturally does a better job of capturing the variation of expected market returns. To explore the dominant variables in constructing the PLS index, Fig. 3 exhibits the top five individual disagreement measures at each point in time when conducting the out-of-sample forecasting. In general, consumption growth forecast dispersion, TBL forecast dispersion, realized personal financial improvement dispersion, business condition forecast dispersion, and house purchase condition forecast dispersion are more likely to be chosen.

In this paper, we measure disagreement with the first PLS factor. One natural question is how many PLS factors we should use in our setting. Following Kelly and Pruitt (2015), we calculate the Bayesian Information Criterion (BIC) via the Krylov representation method and find that only one factor is chosen statistically. To see that this is true, Table A2 reports the R^2s and R^2_{OS}s with the first to sixth moment PLS factors in predicting market returns, where the PLS factors are extracted by using the automatic proxy-selection algorithm in Kelly and Pruitt (2015). The results show that the second to sixth PLS factors do not have any in- and out-of-sample forecasting power, thereby supporting our choice of focusing on the first PLS factor.

In summary, extant disagreement measures do have a common component that is able to predict market returns, and the forecasting power depends on how we aggregate information across individual measures.

3.3. Controlling for economic predictors

This section examines whether the forecasting power of the disagreement index on market returns remains significant after controlling for extant economic predictors. We consider the 14 economic predictors in Welch and Goyal (2008), output gap in Cooper and Priestley (2009), and aggregate short interest in Rapach, Ringgenberg, and Zhou (2016), and run the following regression:

$$R_{t+1} = \alpha + \beta D_t + \psi Z_t + \varepsilon_{t+1}, \tag{13}$$

where Z_t is one of the 16 economic predictors.

Table 4 reports the results. For comparison, the left panel considers the predictive power of the 16 economic predictors and shows that only four variables, the long-term bond return, term spread, output gap, and aggregate short interest, are able to significantly predict market returns. The right panel shows that controlling for extant economic predictors does not reduce the forecasting power of the disagreement index. For example, when controlling for output gap, the corresponding slope slightly decreases to -0.75 in absolute value and is significant at the 1%

level. When the disagreement index and aggregate short interest are jointly used as predictors, the regression slope on the disagreement index remains at a value of -0.84, which is almost the same as without controlling for the aggregate short interest. In the last row, we consider a kitchen sink regression by including all the economic predictors. To handle highly correlated predictors, we estimate the regression slopes with the elastic net method, which has been successfully used in Rapach, Strauss, and Zhou (2013) and Kozak, Nagel, and Santosh (2020) for time series and cross-sectional predictability. The result shows that the forecasting power of the disagreement index remains quantitatively the same as in the case of using the disagreement index alone. Therefore, the predictive ability of the disagreement index is not subsumed by extant economic predictors and it contains independent information beyond these economic predictors.

3.4. Controlling for uncertainty measures

In the literature, disagreement has two alternative interpretations: investor heterogeneity and uncertainty. For example, Anderson, Ghysels, and Juergens (2005) show theoretically and empirically that investor heterogeneity matters for asset pricing and measure it with analyst forecast dispersion. In contrast, Wang, Yan, and Yu (2017) proxy analyst forecast dispersion for uncertainty. While these two alternative explanations can be reconciled by the theory of Atmaz and Basak (2018), it remains empirically interesting to explore whether the disagreement index is different from extant uncertainty measures. Specifically, we employ eight uncertainty measures, including economic uncertainty (Bali, Brown, and Caglayan, 2014), treasury implied volatility (Choi, Mueller, and Vedolin, 2017), financial uncertainty and macro uncertainty (Jurado, Ludvigson, and Ng, 2015), economic policy uncertainty (Baker, Bloom, and Davis, 2016), news implied volatility (Manela and Moreira, 2017), sample variance (Welch and Goyal, 2008), and VIX.

The left panel of Table 5 reports correlations of the disagreement index with the eight uncertainty measures. Consistent with Anderson, Ghysels, and Juergens (2009), the disagreement index does positively correlate with uncertainty except for economic policy uncertainty. For example, the correlation of the disagreement index is 0.33 with treasury implied volatility and 0.24 with macro uncertainty.

Then, we investigate the forecasting power of disagreement by controlling for macro uncertainty as:

$$R_{t+1} = \alpha + \beta D_t + \psi U_t + \varepsilon_{t+1}, \tag{14}$$

where U_t is one of the eight uncertainty measures. As a benchmark, the middle panel of Table 5 shows that extant uncertainty measures cannot significantly predict market returns with one exception, namely financial uncertainty in Jurado, Ludvigson, and Ng (2015). However, the forecasting sign of financial uncertainty seems inconsistent with the asset pricing theory that higher uncertainty implies higher risk premium. The right panel shows that the disagreement index remains significant in predicting market returns after controlling for extant uncertainty measures. For example, in the kitchen sink regression, which includes all the eight uncertainty measures, the slope on the disagreement index is still -0.89, close to the case without any controls (-0.83%). Overall, while the disagreement index is positively correlated with extant uncertainty measures in general, it contains different information for future market returns.

3.5. Economic value with disagreement prediction

In this section, we examine the economic value of forecasting market returns with the disagreement index from the perspective of investing. Following Ferreira and Santa-Clara (2011) and many others, we explore the certainty equivalent return (CER) gain and Sharpe ratio. The higher the CER gain and Sharpe ratio, the larger the risk-rewarded returns when using the disagreement index.

Suppose a mean-variance investor invests her wealth in the stock market and the one-month T-bill. At the start of each month, she allocates a proportion of w_t to the stock market to maximize her next one-month expected utility

$$U(R_p) = \mathrm{E}(R_p) - \frac{\gamma}{2}\mathrm{Var}(R_p), \tag{15}$$

where R_p is the return of the investor's portfolio, $\mathrm{E}(R_p)$ and $\mathrm{Var}(R_p)$ are the mean and variance of the market returns, and γ is the investor's risk aversion.

Let R_{t+1} and $R_{f,t+1}$ be the market return and T-bill rate. The investor's portfolio return at the end of each month is

$$R_{p,t+1} = w_t R_{t+1} + R_{f,t+1}, \tag{16}$$

where $R_{f,t+1}$ is known at t. With a simple calculation, the optimal portfolio weight is

$$w_t = \frac{1}{\gamma}\frac{\hat{R}_{t+1}}{\hat{\sigma}^2_{t+1}}, \tag{17}$$

where $\hat{R}_{t+1}$ and $\hat{\sigma}^2_{t+1}$ are the investor's estimates of the mean and variance of the market returns based on information up to time t.

The CER of the portfolio is

$$\mathrm{CER} = \hat{\mu}_p - \frac{\gamma}{2}\hat{\sigma}^2_p, \tag{18}$$

where $\hat{\mu}_p$ and $\hat{\sigma}^2_p$ are the mean and variance of the investor's portfolio over the out-of-sample evaluation period. The CER can be interpreted as the compensation to the investor for holding the market portfolio. The difference between the CERs for the investor using the predictive regression based on disagreement and the historical return mean is naturally an economic measure of predictability significance.

Table 6 presents the economic value generated by optimally trading on the disagreement index for the investor with a risk aversion of 3 and 5, respectively. That is, we report the CER difference between the strategy using the disagreement forecast and the strategy using the historical return mean. We annualize the CER by multiplying it by 1,200 so that the CER difference denotes the percentage gain per year for the investor to use the disagreement index forecast instead of the historical return mean. Following Campbell and Thompson (2008), we assume that the investor uses a ten-year moving window of past monthly returns to estimate the variance of market returns, and constraints w_t to lie between 0 and 1 to exclude extreme cases.

For comparison, we also consider the alternative PCA and equal-weight disagreement indexes. The results show that among the three disagreement indexes, the PLS disagreement index performs the best and the PCA index performs the worst, which is consistent with the result in Table 3 that both the PLS and equal-weight disagreement indexes can generate significant R^2_{OS}s at the one-month horizon. In Panel A, when there is no transaction cost, the annualized CER gain by using the PLS disagreement index is 2.50%, suggesting that investing with the PLS disagreement index forecast can generate a 2.50% greater risk-adjusted return relative to the historical return mean. The monthly Sharpe ratio is 0.18 and much higher than the market Sharpe ratio of 0.10 in our sample period. When there is a transaction cost of 50 basis points, the CER gain by using the PLS disagreement index is 1.92%, which is still economically sizeable. The corresponding Sharpe ratio is 0.16. Panel B shows similar results when the investor's risk aversion is 5. For example, the CER gain is 2.68% without transaction costs and is 1.88% with a transaction cost of 50 basis points. In summary, the PLS disagreement index is able to deliver considerable economic value for a mean-variance investor.

3.6. Alternative econometric methods

In the previous sections, we have shown that market returns can be significantly predicted by the PLS disagreement index. This section examines whether the result is robust to alternative econometric methods. Particularly, we consider six LASSO-related machine learning methods: equal-weight LASSO, combination LASSO (Han, He, Rapach, and Zhou, 2020), encompassing LASSO (Han, He, Rapach, and Zhou, 2020), adaptive LASSO (Freyberger, Neuhierl, and Weber, 2020), egalitarian LASSO (Diebold and Shin, 2019), and elastic net (Kozak, Nagel, and Santosh, 2020). These six methods are introduced in detail in the Online Appendix.

Table 7 reports the results. There are three observations. First, the out-of-sample R^2_{OS}s are all significant at the one- to 12-month horizons, which confirms the predictability of the PLS disagreement index on market returns. For example, with the elastic net method, the R^2_{OS}s are 1.36% and 8.43% at the one- and 12-month horizons, respectively, and significant at the 5% level. Second, the disagreement index when using the equal-weight LASSO method significantly improves the forecasting power of the equal-weight disagreement index in Section 3.2. The R^2_{OS} increases from 0.90% to 1.26% at the one-month horizon and from 9.41% to 12.08% at the 12-month horizon, thereby suggesting that machine learning techniques are useful for return predictability. Finally, while these six alternative methods work well for predicting market returns, they underperform the PLS. This finding lends empirical support to Kelly and Pruitt's (2015) conclusion that the PLS forecast is asymptotically consistent and will generate the minimum MSFE so long as the consistency condition is satisfied.

To explore which individual disagreement measures are important in predicting market returns, Figs. A1 and A2 in the Online Appendix plot the selected measures and their frequencies according to the six LASSO-related methods at each point in time when conducting out-of-sample forecasting. Over the 1991:02–2018:12 out-of-sample period, some measures are commonly and frequently selected by all the methods. For example, the housing starts forecast dispersion and business condition forecast dispersion are the two most important individual measures and are selected by all six methods with a probability of 100%. The next three important measures are the CPI forecast dispersion, TBL forecast dispersion, and value-weighted analyst forecast dispersion, which are commonly selected with a probability of around 50%. In contrast, the disagreement measures based on the standardized unexplained volume, idiosyncratic volatility, and option open interest are rarely selected by any of the six LASSO-related methods. These results are generally consistent with Fig. 3 and suggest that disagreement measures that are based on professional and household forecasts are equally important in predicting market returns, whereas measures that are based on market information are not.

4. Economic implications

This section shows that the predictability of the disagreement index is consistent with the theory in Atmaz and Basak (2018). In particular, we test four implications. The disagreement index 1) predicts market returns asymmetrically, with stronger power in high-sentiment periods, 2) negatively predicts investors' ex post return forecast errors, 3) predicts market returns via a cash flow channel in the sense of Campbell (1991), and 4) can explain the positive relation between trading volume and market volatility.

4.1. Asymmetric forecasting power

One key implication in Atmaz and Basak (2018) is that disagreement should display an asymmetric forecasting pattern in different market states. The reason is that when investors are relatively pessimistic, the first and second channels have different forecasting signs and are likely to offset each other, making the disagreement-return relation insignificant. In contrast, when investors are overly optimistic, the second channel dominates the first channel and hence disagreement negatively predicts future stock returns. In the following, we use the investor sentiment index of Baker and Wurgler (2006) to test whether the forecasting power of the disagreement index is asymmetric over the high- and low-sentiment periods, where a month is defined as high if the past 18-month average sentiment index is positive, and low otherwise.

4.1.1. Time series evidence

Following Rapach, Strauss, and Zhou (2010), we calculate the in-sample R^2s in high- and low-sentiment periods as:

$$R_c^2 = 1 - \frac{\sum_{t=1}^{T} S_t^c (\hat{\varepsilon}_t)^2}{\sum_{t=1}^{T} S_t^c (R_t - \bar{R})^2}, \quad c = \text{high, low}, \tag{19}$$

where S_t^{high} (S_t^{low}) is an indicator that takes a value of one when month t is in a high- (low-) sentiment period and zero otherwise, $\hat{\varepsilon}_t$ is the fitted residual based on the in-sample estimate, $\bar{R}$ is the full sample mean of R_t, and T is the number of observations for the full sample. Note that, unlike the full-sample R^2 statistic, the R^2_{high} and R^2_{low} statistics can be either positive or negative. Similarly, we can also calculate the R^2_{OS} in high- and low- sentiment periods separately. Another way to test the forecasting asymmetry is to run the following state-dependent regression:

$$R_{t+1} = \alpha + \beta_{\text{high}} S_t^{\text{high}} D_t + \beta_{\text{low}} S_t^{\text{low}} D_t + \varepsilon_{t+1}. \tag{20}$$

Table 8 shows that the forecasting power of the disagreement index is concentrated in high-sentiment periods. In Panel A, the R^2 and R^2_{OS} are 5.28% and 3.69% in high-sentiment periods, and 0.80% and −0.55% in low-sentiment periods, respectively. In Panel B, the regression slope of the disagreement index in high-sentiment periods is −1.12 with a t-value of −4.71, but it is only −0.42 with an insignificant t-value of −1.21 in low-sentiment periods. Therefore, the predictability of disagreement on market returns is asymmetric and concentrated in high-sentiment periods.

4.1.2. Cross-sectional evidence

In a multiple stock economy, Atmaz and Basak (2018) suggest that the forecasting power of disagreement should be asymmetric across stocks, stronger among stocks with optimistic investor expectation and weaker or insignificant among stocks with pessimistic investor expectation. In contrast to Miller's (1977) results, this implication holds even in the absence of short-sale constraints, so long as there are infinite risk-averse investors. For this reason, we test the implication based on portfolios sorted by expectation directly.

Following Bordalo, Gennaioli, La Porta, and Shleifer (2019), we proxy the analyst LTG forecast for investor expectation at the firm level and construct ten decile portfolios at the end of December each year. The portfolios are subsequently held for one year. In the 1982–2018 period, the portfolio with low-LTG forecast earns an annual return of 13.58% and the portfolio with high-LTG forecast earns an annual return of 7.89%, with the difference between the high- and low-LTG forecast portfolios equal to 5.69% per year. Panel A of Fig. 4 plots the regression slopes of predicting the ten decile portfolio returns with the disagreement index. Apparently, the slope increases in magnitude from -0.61 for the portfolio with low-LTG forecast to -1.09% for the portfolio with high-LTG forecast.

Also, to explore the time-varying effect of the average bias of investor expectation, we run the following state-dependent regression:

$$R_{i,t+1} = \alpha_i + \beta_{\text{high},i} S_t^{\text{high}} D_t + \beta_{\text{low},i} S_t^{\text{low}} D_t + \varepsilon_{i,t+1}. \tag{21}$$

Panels B and C of Fig. 4 plot the regression slopes in high- and low-sentiment periods, respectively. As expected, the forecasting power of the disagreement index is concentrated in high-sentiment periods. β_{high} monotonically increases in magnitude from -0.57 for the low-LTG forecast portfolio to -1.96 for the high-LTG forecast portfolio. In contrast, β_{low} is flat and displays a slightly upward trend. Also, Fig. A3 in the Online Appendix shows that portfolios with low institutional ownership, high beta, or high IVOL earn lower average returns among high-disagreement periods and confirms the argument that disagreement and arbitrage costs have an interaction effect (see, e.g., Hong and Sraer, 2016). In general, the predictability of disagreement is both time series and cross-sectionally asymmetric and is stronger among stocks with optimistic cash flow expectation in high sentiment periods.

4.2. Disagreement and expectation of market returns

In the previous section we have linked disagreement with investor expectation (measured by investor sentiment) in an indirect manner. In this section we examine the relation of disagreement with investor expectation of market returns directly. According to Atmaz and Basak (2018), since disagreement amplifies investor optimism, it should be negatively related to ex post return forecast errors.[4]

Specifically, we consider four measures of investor expectation of 12-month-ahead market returns. The first measure is the value-weighted aggregate analysts' return forecast, where the analysts' return forecast of an individual stock is defined as the mean of 12-month-ahead analysts' target prices divided by current price (Engelberg, McLean, and Pontiff, 2020), and the target prices are restricted to those reported in the past one month. The second measure is the University of Michigan Surveys of Consumers Attitudes. Following Das, Kuhnen, and Nagel (2019), we use responses to the question "Suppose that tomorrow someone were to invest one thousand dollars in a type of mutual fund known as a diversified stock fund. What do you think is the percent chance that this one thousand dollar investment will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?" The third and fourth measures are Graham-Harvey's survey of CFOs and Robert Shiller's survey of individual investor confidence in the stock market, which are constructed strictly following Greenwood and Shleifer (2014). The first three expectation measures have the same measurement unit as the realized market returns, whereas the last one, Shiller's survey, is based on binary variables. As such, when calculating the ex post return forecast errors, we project the aggregate analysts' return forecast on Shiller's survey, so that the projected time series has the same measurement unit as the realized returns.

Table 9 reports the results. In Panel A, the disagreement index is positively related to the four investor expectation measures, with significant correlations ranging from 0.24 to 0.35. In terms of economic magnitude, a one-standard-deviation increase in disagreement is associated with 3.26%, 2.71%, 0.57%, and 2.16% increases in investor expectation of 12-month-ahead market returns with the analyst's return forecast, University of Michigan survey, Graham-Harvey's survey, and Shiller's survey, respectively. Untabulated results also confirm Greenwood and Shleifer (2014)'s conclusion that these investor expectation measures negatively predict 12-month-ahead market returns. In Panel B, when regressing ex post return forecast errors on the disagreement index, we find that

[4]We thank the referee for this and many other intriguing suggestions

all the regression coefficients are significantly negative. For example, a one-standard-deviation increase in disagreement is associated with a 7.34% increase in the analysts' return forecast error (i.e., the deviation of analysts' return forecast from the realized return increases by 7.34%), and the disagreement index explains about one quarter of the variations of analysts' return forecast errors (i.e., $R^2 = 23.26\%$). Overall, Table 9 suggests that disagreement is closely linked with investor expectation of market returns for both sophisticated and retailed investors.

4.3. Relations of disagreement with cash flow news and discount rate news

This section examines the contemporaneous relations of disagreement with cash flow news and discount rate news, so that we can disentangle the forecasting channel in the sense of Campbell (1991).

Following Campbell (1991), the log total market return can be decomposed into three components,

$$\tilde{R}_t \approx \mathrm{E}_{t-1}(R_t) + \mathrm{CF}_t - \mathrm{DR}_t, \tag{22}$$

where CF_t and DR_t are cash flow news and discount rate news, and they are defined as

$$\mathrm{CF}_t = (\mathrm{E}_t - \mathrm{E}_{t-1}) \sum_{j=0}^{\infty} \kappa^j \Delta d_{t+j} = (\Delta d_t - \mathrm{E}_{t-1} \Delta d_t) + (\mathrm{E}_t - \mathrm{E}_{t-1}) \sum_{j=1}^{\infty} \kappa^j \Delta d_{t+j}, \tag{23}$$

$$\mathrm{DR}_t = (\mathrm{E}_t - \mathrm{E}_{t-1}) \sum_{j=1}^{\infty} \kappa^j \tilde{R}_{t+j}, \tag{24}$$

where Δd_{t+j} and $\tilde{R}_{t+j}$ are the log dividend growth and log total market return at time $t + j$, and κ is a log-linearization constant slightly less than one. In other words, CF_t and DR_t are return innovations due to updates in expectations of current and future cash flows and future expected returns, respectively.

Atmaz and Basak (2018) posit that, after positive cash flow news, say $\Delta d_t - \mathrm{E}_{t-1} \Delta d_t > 0$, investors whose beliefs are supported by the cash flow news become relatively wealthier, which makes them more optimistic about future cash flows or discount rates or both, and, consequently, increases disagreement. In contrast, after negative cash flow news, investors who have been optimistic become relatively poorer and pessimistic, thereby shrinking disagreement. This argument suggests that both CF_t and DR_t can positively affect disagreement and drive its movements.

To explore which component is the main driver of disagreement, we use contemporaneous CF_t and DR_t as the targets in Eq. (10) to extract a cash flow news-based PLS disagreement index and a discount rate news-based PLS disagreement index, and then examine their power in predicting future market returns. The results are reported in Table 10, where the cash flow news and discount rate news are estimated based on individual VARs comprising the total market return, dividend-price ratio, and one of the rest of the 15 economic predictors explored in Table 4. We always include the dividend-price ratio in the VARs because Engsted, Pedersen, and Tanggaard (2012) show that it is important to include this variable to properly estimate the cash flow and discount news components. In the last row of Table 10, we also consider the decomposition based on a VAR comprising the total market return, log dividend-price ratio, and the first three principal components extracted from the 15 economic predictors.

Table 10 shows that only the cash flow news-based disagreement index has forecasting power on market returns. For example, when the cash flow news and discount rate news are estimated with the VAR comprising the total market return and dividend-price ratio, a one-standard-deviation increase in the cash flow news-based disagreement index predicts a 0.65% decrease in the next one-month market return, while the discount rate news-based disagreement index displays no power. This finding echoes Section 3.2, in which we show statistically that there is only one PLS factor exhibiting forecasting power on market returns. Thus, we conclude that the ability of disagreement to predict market returns is more likely to operate via a cash flow channel in the sense of Campbell (1991).

4.4. *Relations of disagreement with trading volume and market volatility*

In Atmaz and Basak (2018), in the absence of disagreement, trading volume is zero and market volatility is constant. In the presence of disagreement, however, higher disagreement leads to both higher trading volume and higher market volatility, thereby suggesting that disagreement is the driver of the positive volume-volatility relationship.

To test the implication, we estimate the volume-volatility elasticity in month t as the slope of regressing the daily change in market turnover on the daily change in volatility within month t, and then regress the monthly elasticity on the lagged disagreement index. For robustness, we consider four daily volatility measures, including realized volatility, realized semi-volatility, and median realized volatility based on the S&P 500 index returns from five-minute intervals from Andersen, Dobrev, and Schaumburg (2012), and realized volatility of the S&P 500 index futures contract returns from five-minute intervals from Johnson (2019).

Panel A of Table 11 shows that the disagreement index positively predicts the volume-volatility elasticity. The intuition is that increased disagreement increases the investor expectation bias, which in turn increases both the fluctuation of stock price and the trading demand (due to the increased weight of investors with relatively different beliefs), thereby increasing the volume-volatility elasticity. In a more intuitive way, we find in Table A3 that the disagreement index positively predicts the correlation between trading volume and market volatility. For example, a one-standard-deviation increase in disagreement predicts a 5.22% increase in the volume-volatility correlation of next month when market volatility is estimated with the realized volatility.

To corroborate Panel A, we decompose market volatility into two components: one is contemporaneously related to disagreement and extracted via the PLS method and the other is unrelated to disagreement. Then, we regress the one-month-ahead trading volume on these two volatility components and report the results in Panel B of Table 11. As expected, the disagreement-related volatility significantly positively predicts future trading volume, whereas the disagreement-unrelated volatility does not have any predictive power. Similarly, when decomposing trading volume into disagreement-related and -unrelated components, we find that the disagreement-related volume predicts future market volatility but the disagreement-unrelated volume fails to do so, which is reported in Panel B of Table A3.

In sum, this section provides empirical support to Atmaz and Basak's (2018) implication that disagreement seems to be a key driver of the positive volume-volatility relation.

5. Conclusion

This paper examines whether extant individual disagreement measures are agreeable and proposes a disagreement index by using the PLS methodology in Kelly and Pruitt (2013, 2015). We show that this PLS disagreement index significantly predicts market returns both in- and out-of-sample. Consistent with the theory in Atmaz and Basak (2018), the disagreement index asymmetrically predicts market returns with greater power in high-sentiment periods, is negatively related to investors' ex post return forecast errors, predicts market returns through a cash flow channel, and is able to explain the positive volume-volatility relation.

There are some open issues for future research. First, it would be interesting to apply the disagreement index to other markets, such as bonds, commodities, and currencies, to see whether the forecasting power remains significant. Second, it would be of interest to construct aggregate disagreement indexes at different frequencies, such as daily or weekly, so that investors could use them for real time investing. Finally, as Hong and Stein (2007) posit that there are two main sources of disagreement, differences in information sets and differences in models that investors use to interpret information, it would be interesting to disentangle them.

References

Ajinkya, B. B., Atiase, R. K., Gift, M. J., 1991. Volume of trading and the dispersion in financial analysts' earnings forecasts. Accounting Review 66, 389–401.

Andersen, T., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169, 75–93.

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

Anderson, E. W., Ghysels, E., Juergens, J. L., 2009. The impact of risk and uncertainty on expected returns. Journal of Financial Economics 94, 233–263.

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

Atmaz, A., Basak, S., 2018. Belief dispersion in the stock market. Journal of Finance 73, 1225–1279.

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

Baker, S. R., Bloom, N., Davis, S. J., 2016. Measuring economic policy uncertainty. Quarterly Journal of Economics 131, 1593–1636.

Bali, T. G., Brown, S. J., Caglayan, M. O., 2014. Macroeconomic risk and hedge fund returns. Journal of Financial Economics 114, 1–19.

Barber, B. M., Odean, T., 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.

Berkman, H., Dimitrov, V., Jain, P. C., Koch, P. D., Tice, S., 2009. Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements. Journal of Financial Economics 92, 376–399.

Boehme, R. D., Danielsen, B. R., Sorescu, S. M., 2006. Short-sale constraints, differences of opinion, and overvaluation. Journal of Financial and Quantitative Analysis 41, 455–487.

Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2019. Diagnostic expectations and stock returns. Journal of Finance 74, 2839–2874.

Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Over-reaction in macroeconomic expectations. American Economic Review, forthcoming.

Campbell, J. Y., 1991. A variance decomposition for stock returns. Economic Journal 101, 157–179.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Carlin, B. I., Longstaff, F. A., Matoba, K., 2014. Disagreement and asset prices. Journal of Financial Economics 114, 226–238.

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

Chen, Y., Eaton, G. W., Paye, B. S., 2018. Micro (structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity. Journal of Financial Economics 130, 48–73.

Chinco, A., Clark-Joseph, A. D., Ye, M., 2019. Sparse signals in the cross-section of returns. Journal of Finance 74, 449–492.

Choi, H., Mueller, P., Vedolin, A., 2017. Bond variance risk premiums. Review of Finance 21, 987–1022.

Clark, T. E., West, K. D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

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

Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. Review of Financial Studies 22, 2801–2833.

Das, S., Kuhnen, C. M., Nagel, S., 2019. Socioeconomic status and macroeconomic expectations. Review of Financial Studies 33, 395–432.

DeVault, L., Sias, R. W., Starks, L. T., 2019. Sentiment metrics and investor demand. Journal of Finance 74, 985–1024.

Diebold, F. X., Shin, M., 2019. Machine learning for regularized survey forecast combination: Partially-egalitarian LASSO and its derivatives. International Journal of Forecasting 35, 1679–1691.

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

Edelen, R. M., Ince, O. S., Kadlec, G. B., 2016. Institutional investors and stock return anomalies. Journal of Financial Economics 119, 472–488.

Engelberg, J., McLean, R. D., Pontiff, J., 2020. Analysts and anomalies. Journal of Accounting and Economics 69, 1–13.

Engsted, T., Pedersen, T. Q., Tanggaard, C., 2012. Pitfalls in VAR based return decompositions: A clarification. Journal of Banking & Finance 36, 1255–1265.

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

Ferreira, M. A., Santa-Clara, P., 2011. Forecasting stock market returns: The sum of the parts is more than the whole. Journal of Financial Economics 100, 514–537.

Freyberger, J., Neuhierl, A., Weber, M., 2020. Dissecting characteristics nonparametrically. Review of Financial Studies 33, 2326–2377.

Garfinkel, J. A., 2009. Measuring investors' opinion divergence. Journal of Accounting Research 47, 1317–1348.

Ge, L., Lin, T.-C., Pearson, N. D., 2016. Why does the option to stock volume ratio predict stock returns? Journal of Financial Economics 120, 601–622.

Giglio, S., Kelly, B., Pruitt, S., 2016. Systemic risk and the macroeconomy: An empirical evaluation. Journal of Financial Economics 119, 457–471.

Golez, B., Koudijs, P., 2018. Four centuries of return predictability. Journal of Financial Economics 127, 248–263.

Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. Review of Financial Studies 27, 714–746.

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

Hamilton, J. D., 2018. Why you should never use the Hodrick-Prescott filter. Review of Economics and Statistics 100, 831–843.

Han, Y., He, A., Rapach, D., Zhou, G., 2020. Firm characteristics and expected stock returns. Unpublished working paper. Washington University in St. Louis.

Hong, H., Sraer, D. A., 2016. Speculative betas. Journal of Finance 71, 2095–2144.

Hong, H., Stein, J. C., 2007. Disagreement and the stock market. Journal of Economic Perspectives 21, 109–128.

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

Huang, D., Kilic, M., 2019. Gold, platinum, and expected stock returns. Journal of Financial Economics 132, 50–75.

Jacobsen, B., Marshall, B. R., Visaltanachoti, N., 2019. Stock market predictability and industrial metal returns. Management Science 65, 2947–3448.

Jiang, H., Sun, Z., 2014. Dispersion in beliefs among active mutual funds and the cross-section of stock returns. Journal of Financial Economics 114, 341–365.

Johnson, T. L., 12 2019. A fresh look at return predictability using a more efficient estimator. Review of Asset Pricing Studies 9, 1–46.

Jondeau, E., Zhang, Q., Zhu, X., 2019. Average skewness matters! Journal of Financial Economics 134, 29–47.

Jurado, K., Ludvigson, S. C., Ng, S., 2015. Measuring uncertainty. American Economic Review 105, 1177–1216.

Kelly, B., Pruitt, S., 2013. Market expectations in the cross-section of present values. Journal of Finance 68, 1721–1756.

Kelly, B., Pruitt, S., 2015. The three-pass regression filter: A new approach to forecasting using many predictors. Journal of Econometrics 186, 294–316.

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross section. Journal of Financial Economics 135, 271–292.

Lee, J. D., Sun, D. L., Sun, Y., Taylor, J. E., 2016. Exact post-selection inference, with application to the LASSO. Annals of Statistics 44, 907–927.

Li, D., Li, G., 2017. Household investors' belief dispersion, trading volume, and excess returns. Unpublished working paper. Federal Reserve Board.

Li, F. W., 2016. Macro disagreement and the cross-section of stock returns. Review of Asset Pricing Studies 6, 1–45.

Li, J., Yu, J., 2012. Investor attention, psychological anchors, and stock return predictability. Journal of Financial Economics 104, 401–419.

Li, Y., Ng, D. T., Swaminathan, B., 2013. Predicting market returns using aggregate implied cost of capital. Journal of Financial Economics 110, 419–436.

Light, N., Maslov, D., Rytchkov, O., 2017. Aggregation of information about the cross section of stock returns: A latent variable approach. Review of Financial Studies 30, 1339–1381.

Lyle, M. R., Wang, C. C., 2015. The cross section of expected holding period returns and their dynamics: A present value approach. Journal of Financial Economics 116, 505–525.

Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. Journal of Financial Economics 123, 137–162.

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

Møller, S. V., Rangvid, J., 2015. End-of-the-year economic growth and time-varying expected returns. Journal of Financial Economics 115, 136–154.

Pyun, S., 2019. Variance risk in aggregate stock returns and time-varying return predictability. Journal of Financial Economics 132, 150–174.

Rapach, D. E., Ringgenberg, M. C., Zhou, G., 2016. Short interest and aggregate stock returns. Journal of Financial Economics 121, 46–65.

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

Rapach, D. E., Strauss, J. K., Zhou, G., 2013. International stock return predictability: What is the role of the United States? Journal of Finance 68, 1633–1662.

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

Tibshirani, R. J., Taylor, J., Lockhart, R., Tibshirani, R., 2016. Exact post-selection inference for sequential regression procedures. Journal of the American Statistical Association 111, 600–620.

Wang, H., Yan, J., Yu, J., 2017. Reference-dependent preferences and the risk-return trade-off. Journal of Financial Economics 123, 395–414.

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

Wold, H., 1966. Estimation of rincipal components and related models by iterative least squares. In P. R. Krishnajah (Ed.), Multivariate Analysis (pp. 391–420). Academic Press, New York.

Yu, J., 2011. Disagreement and return predictability of stock portfolios. Journal of Financial Economics 99, 162–183.

Yu, J., Yuan, Y., 2011. Investor sentiment and the mean–variance relation. Journal of Financial Economics 100, 367–381.

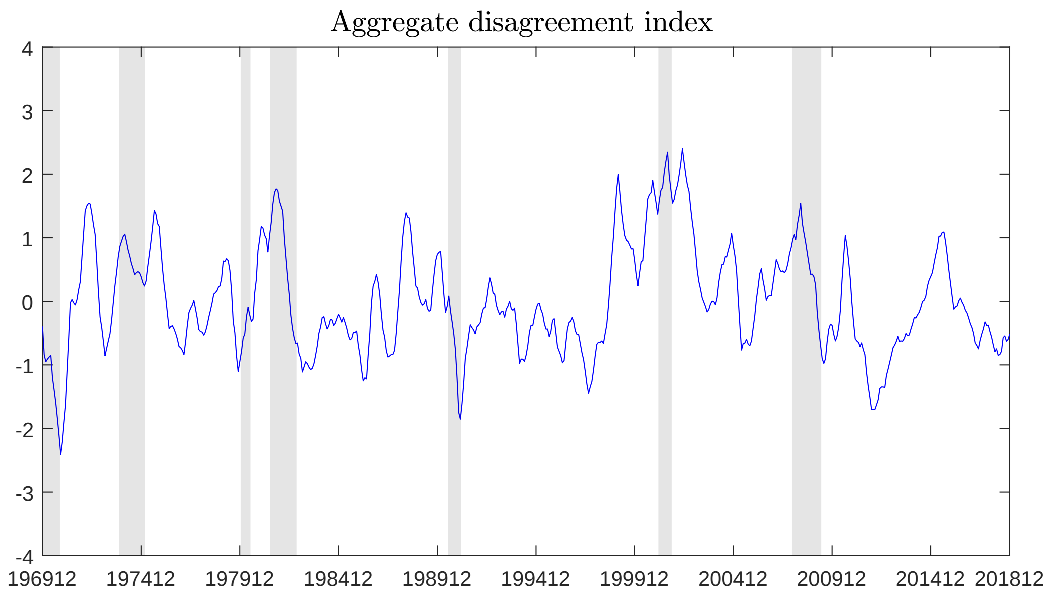

Fig. 1. This figure plots the time series dynamics of the disagreement index constructed using the PLS method in Kelly and Pruitt (2013, 2015). Grey shadow bars denote NBER recessions. The sample period is 1969:12–2018:12.

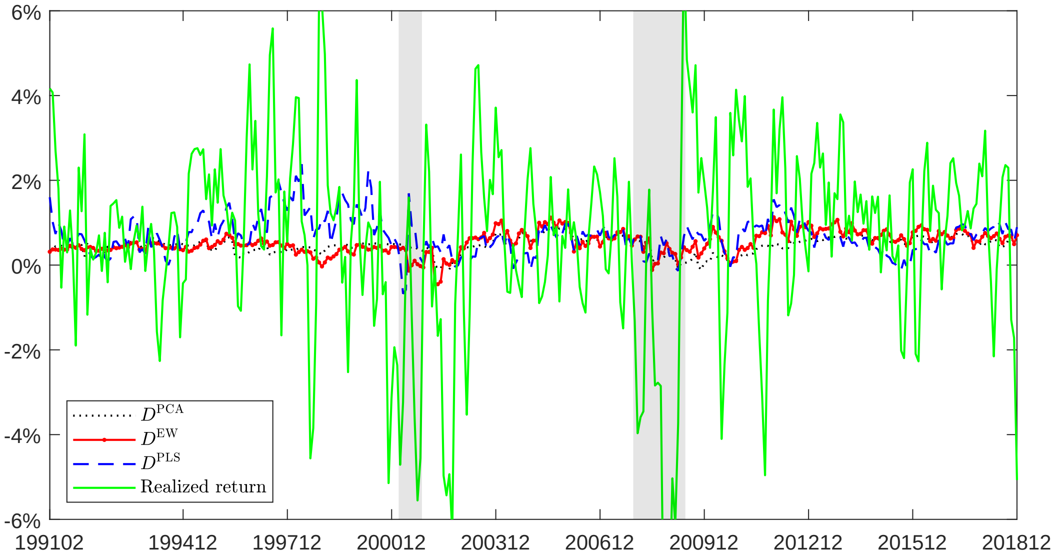

Fig. 2. This figure plots the out-of-sample three-month market return forecasts with the PCA, equal-weight, and PLS disagreement indexes, respectively. For comparison, the figure also plots the realized three-month market returns. Grey shadow bars denote NBER recessions. The out-of-sample period is 1991:02–2018:12.

Fig. 3. This figure plots the top five individual disagreement measures in the PLS disagreement index at each point in time when conducting out-of-sample forecasting.

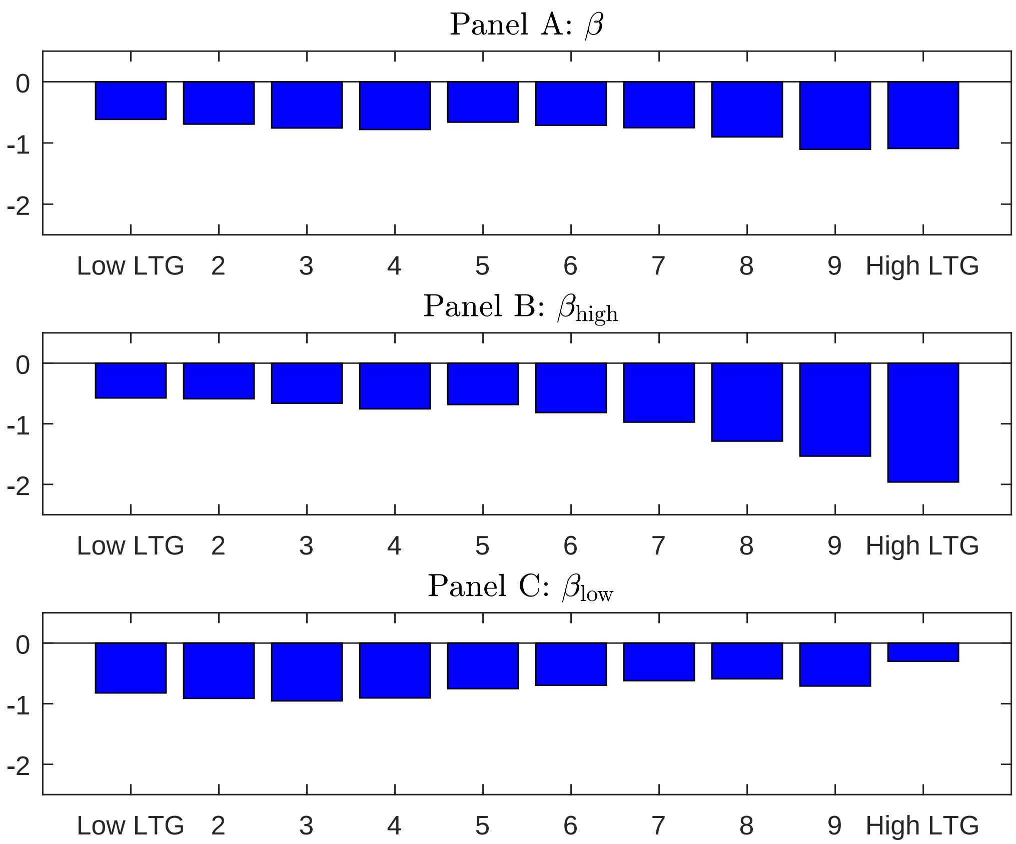

Fig. 4. This figure plots the regression slopes of predicting portfolio excess returns sorted by analyst long-term growth rate (LTG) forecast with the disagreement index as

$$R_{i,t+1} = \alpha_i + \beta_i D_t + \varepsilon_{i,t+1}$$

for Panel A, and

$$R_{i,t+1} = \alpha_i + \beta_{\text{high},i} S_t^{\text{high}} D_t + \beta_{\text{low},i} S_t^{\text{low}} D_t + \varepsilon_{i,t+1}$$

for Panels B and C. Low- (high-) LTG refers to the portfolio with the low (high) analyst LTG forecast and is constructed as in Bordalo, Gennaioli, La Porta, and Shleifer (2019). S_t^{high} (S_t^{low}) is a dummy variable that equals one if month t is in high- (low-) sentiment periods and zero if month t is in low- (high-) sentiment periods (Baker and Wurgler, 2006). The sample period is 1982:01–2018:12.

Table 1
Summary statistics of individual disagreement measures.

This table reports the summary statistics of 24 individual disagreement measures used in this paper. The first 13 measures are obtained from the Survey of Professional Forecasters (SPF) at a quarterly frequency; each is defined by the level or growth difference between the 75th and 25th percentiles of the forecasts. D^{Yu} and D^{HS} are value- and beta-weighted analyst forecast dispersions (Yu, 2011; Hong and Sraer, 2016). The next six are household belief dispersions on macroeconomic conditions from the University of Michigan Surveys of Consumers Attitudes. D^{SUV} is a disagreement measure based on the standardized unexplained trading volume of NYSE stocks (Garfinkel, 2009). D^{IVOL} is the value-weighted idiosyncratic volatility proposed by Boehme, Danielsen, and Sorescu (2006) for measuring investor disagreement. D^{OID} is a disagreement measure defined by the open interest difference of OEX call and put options (Ge, Lin, and Pearson, 2016).

Disagreement measure	Sample Period	Obs	Avg	Std	Min	Max	Skew	Kurt
GDP forecast dispersion (D^{GDP})	1968Q4–2018Q4	201	61.32	41.13	6.80	248.60	1.28	2.53
GDP growth forecast dispersion (D^{GDPg})	1968Q4–2018Q4	201	1.65	0.71	0.71	4.25	1.07	0.81
Industrial production forecast dispersion (D^{IP})	1968Q4–2018Q4	201	1.97	1.03	0.52	6.10	1.14	1.30
Industrial production growth forecast dispersion (D^{IPg})	1968Q4–2018Q4	201	2.73	1.43	0.84	8.04	1.14	0.99
Consumption forecast dispersion (D^{CON})	1968Q4–2018Q4	150	31.61	18.31	5.00	101.87	1.08	1.91
Consumption growth forecast dispersion (D^{CONg})	1968Q4–2018Q4	150	0.97	0.40	0.39	2.79	1.33	2.07
Investment forecast dispersion (D^{INV})	1981Q3–2018Q4	150	22.46	12.47	3.40	57.92	0.48	−0.39
Investment growth forecast dispersion (D^{INVg})	1981Q3–2018Q4	150	3.63	1.19	1.43	8.62	0.71	1.33
Housing starts forecast dispersion (D^{HSG})	1968Q4–2018Q4	201	0.12	0.04	0.05	0.27	0.90	0.53
Housing starts growth forecast dispersion (D^{HSGg})	1968Q4–2018Q4	201	18.70	10.00	6.46	57.34	1.38	1.52
Unemployment rate forecast dispersion (D^{UEP})	1968Q4–2018Q4	201	0.32	0.13	0.15	1.04	1.75	5.10
CPI forecast dispersion (D^{CPI})	1981Q3–2018Q4	150	0.83	0.30	0.38	2.02	1.35	1.89
TBL forecast dispersion (D^{TBL})	1981Q3–2018Q4	150	0.46	0.36	0.04	2.96	3.36	17.15
Value-weighted analyst forecast dispersion (D^{Yu})	1981:12–2018:12	445	3.67	0.61	2.64	5.79	1.04	0.60
Beta-weighted analyst forecast dispersion (D^{HS})	1981:12–2018:12	445	5.15	1.29	3.41	9.62	1.39	1.87
Realized personal financial improvement dispersion (D^{RPF})	1978:01–2018:12	492	−0.44	0.02	−0.50	−0.39	−0.42	−0.37
Expected personal financial improvement forecast dispersion (D^{EPF})	1978:01–2018:12	492	−0.64	0.05	−0.80	−0.50	−0.22	0.24
Business condition forecast dispersion (D^{BC})	1978:01–2018:12	492	−0.42	0.07	−0.69	−0.28	−0.96	1.04
Unemployment condition forecast dispersion (D^{UC})	1978:01–2018:12	492	−0.63	0.08	−0.95	−0.47	−0.60	0.14
Interest rate condition forecast dispersion (D^{IRC})	1978:01–2018:12	492	−0.53	0.08	−0.77	−0.35	−0.17	−0.59
House purchase condition forecast dispersion (D^{HOM})	1978:01–2018:12	492	−0.59	0.08	−0.80	−0.42	−0.02	−0.64
Standardized unexplained volume (D^{SUV})	1968:12–2018:12	589	0.14	1.25	−3.45	3.17	−0.15	−0.68
Idiosyncratic volatility (D^{IVOL})	1968:12–2018:12	589	0.02	0.00	0.01	0.03	1.72	3.25
OEX call/put open interest difference (D^{OID})	1984:02–2018:12	419	0.86	0.09	0.55	1.00	−0.99	0.73

Table 2
Forecasting market returns with individual disagreement measures.

This table presents the regression slope, Newey-West t-value, in-sample R^2, and out-of-sample R^2_{OS} of predicting market returns with individual disagreement measures:

$$R_{t,t+h} = \alpha + \beta D_t + \varepsilon_{t,t+h},$$

where $R_{t,t+h}$ is the cumulative market return between months t and $t+h$ ($h = 1, 3$, or 12), and D_t is one of the 24 individual disagreement measures. The in-sample period is 1969:12–2018:12 and the out-of-sample period is 1991:02–2018:12 (because the accurate release dates of the University of Michigan Surveys of Consumers Attitudes are only available as of January 1991). Statistical significance for R^2_{OS} is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	$h=1$				$h=3$				$h=12$			
Disagreement	β	t-value	R^2	R^2_{OS}	β	t-value	R^2	R^2_{OS}	β	t-value	R^2	R^2_{OS}
D^{GDP}	−0.11	−0.50	0.06	−1.88	−0.22	−1.35	0.72	−5.40	−0.36***	−3.54	7.43	−13.61
D^{GDPg}	−0.27	−1.53	0.38	−3.44	−0.24	−1.62	0.89	−7.18	−0.24**	−2.56	3.21	−16.47
D^{IP}	−0.06	−0.24	0.02	−2.38	−0.07	−0.42	0.08	−5.04	−0.04	−0.31	0.09	−12.18
D^{IPg}	0.04	0.20	0.01	−2.23	−0.21	−1.46	0.65	−9.50	−0.14	−1.09	1.13	−20.03
D^{CON}	−0.11	−0.49	0.07	−1.79	−0.24	−1.21	0.88	−4.34	−0.17	−1.27	1.72	−15.22
D^{CONg}	−0.10	−0.41	0.05	−2.38	−0.19	−0.92	0.59	−5.73	−0.17	−1.29	1.73	−19.59
D^{INV}	−0.24	−1.36	0.31	−2.73	−0.27*	−1.75	1.12	−8.11	−0.14	−0.88	1.15	−12.17
D^{INVg}	0.19	1.31	0.21	−1.60	0.02	0.17	0.01	−4.60	0.02	0.14	0.04	−8.25
D^{HSG}	−0.40**	−2.11	0.85	−5.39	−0.26*	−1.70	1.04	−12.29	−0.21	−1.57	2.48	−23.20
D^{HSGg}	−0.21	−0.99	0.23	−6.49	−0.34**	−2.08	1.70	−24.59	−0.33*	−1.90	5.73	−28.78
D^{UEP}	0.16	0.76	0.13	−0.73	0.13	0.82	0.27	−2.64	0.16	1.47	1.48	−3.38
D^{CPI}	−0.36*	−1.90	0.73	−6.39	−0.27**	−2.26	1.18	−27.39	−0.11	−1.32	0.74	−19.33
D^{TBL}	−0.60**	−2.23	1.94	−4.21	−0.48**	−2.12	3.60	−9.36	−0.25	−1.53	3.84	−13.52
D^{Yu}	−0.30	−1.23	0.35	−2.72	−0.30	−1.32	1.05	−4.07	−0.27	−1.25	3.06	−25.93
D^{HS}	−0.12	−0.42	0.06	−3.67	−0.16	−0.55	0.30	−7.92	−0.23	−0.85	2.52	−11.16
D^{RPF}	−0.22	−1.26	0.26	−3.35	−0.09	−0.55	0.11	−6.13	−0.16	−1.57	1.52	−17.10
D^{EPF}	−0.22	−1.03	0.26	−4.49	−0.16	−1.11	0.40	−9.37	−0.05	−0.49	0.17	−17.36
D^{BC}	−0.44**	−2.48	1.05	−5.77	−0.23	−1.56	0.86	−10.44	−0.08	−0.66	0.37	−19.99
D^{UC}	−0.06	−0.27	0.02	−3.40	−0.03	−0.19	0.02	−6.35	−0.05	−0.37	0.17	−19.27
D^{IRC}	−0.17	−0.72	0.16	−2.64	−0.41**	−2.12	2.66	−12.05	−0.44**	−2.52	11.72	−25.38
D^{HOM}	0.11	0.65	0.07	−0.86	0.01	0.04	0.00	−3.71	−0.19	−1.12	1.94	−20.53
D^{SUV}	−0.30	−1.24	0.23	−2.80	−0.34	−1.48	0.86	−6.86	−0.43*	−1.86	4.81	−18.40
D^{IVOL}	−0.03	−0.15	0.00	−5.46	−0.02	−0.07	0.00	−14.61	0.00	−0.02	0.00	−14.82
D^{OID}	−0.20	−0.73	0.10	−2.73	−0.15	−0.57	0.15	−6.09	−0.24	−0.93	1.47	−15.53

Table 3
Forecasting market returns with aggregate disagreement indexes.

This table presents the regression slope, Newey-West t-value, in-sample R^2, and out-of-sample R^2_{OS} of predicting market returns with disagreement as

$$R_{t,t+h} = \alpha + \beta D_t + \varepsilon_{t,t+h},$$

where $R_{t,t+h}$ is the cumulative market return between months t and $t+h$ ($h = 1, 3$, and 12), and D_t is the PCA, equal-weight (individual measures), or PLS disagreement index. Statistical significance for R^2_{OS} is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. The in- and out-of-sample periods are 1969:12–2018:12 and 1991:02–2018:12, respectively. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Method	β	t-value	R^2	R^2_{OS}
Panel A: $h = 1$				
PCA	−0.38**	−1.96	0.61	0.20
Equal-weight	−0.60***	−2.87	1.46	0.90**
PLS	−0.83***	−3.96	2.52	1.56**
Panel B: $h = 3$				
PCA	−1.13**	−2.14	1.74	1.71***
Equal-weight	−1.73***	−2.99	3.88	3.74***
PLS	−2.24***	−3.82	5.98	7.68***
Panel C: $h = 12$				
PCA	−2.92**	−1.99	2.78	3.04***
Equal-weight	−4.93***	−3.06	7.49	9.41***
PLS	−7.04***	−4.16	13.88	13.26***

Table 4
Controlling for economic variables.

Panel A presents the results of predicting market returns as

$$R_{t+1} = \alpha + \psi Z_t + \varepsilon_{t+1},$$

where Z_t is one of the 14 economic predictors in Welch and Goyal (2008), output gap in Cooper and Priestley (2009), or aggregate short interest in Rapach, Ringgenberg, and Zhou (2016). Panel B reports the results of forecasting market returns with the disagreement index and one economic predictor as

$$R_{t+1} = \alpha + \beta D_t + \psi Z_t + \varepsilon_{t+1}.$$

The last row reports the slope of the disagreement index from an elastic net regression by including all the economic predictors, where the t-value is calculated following Tibshirani, Taylor, Lockhart, and Tibshirani (2016) and Lee, Sun, Sun, and Taylor (2016). The sample period is 1969:12–2018:12. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	Univariate		Bivariate		
Economic predictor	ψ	R^2	β	ψ	R^2
Dividend-price ratio (DP)	0.15	0.11	−0.82***	0.02	2.52
Dividend yield (DY)	0.17	0.15	−0.82***	0.03	2.52
Earning-price ratio (EP)	0.08	0.03	−0.83***	−0.01	2.52
Dividend payout ratio (DE)	0.08	0.03	−0.82***	0.03	2.53
Sample variance (SVAR)	−0.22	0.26	−0.81***	−0.07	2.54
Book-to-market ratio (BM)	0.00	0.00	−0.83***	−0.03	2.53
Net equity expansion (NTIS)	−0.06	0.02	−0.83***	−0.07	2.55
T-bill rate (TBL)	−0.26	0.36	−0.81***	−0.21	2.76
Long-term bond yield (LTY)	−0.15	0.11	−0.82***	−0.12	2.60
Long-term bond return (LTR)	0.42**	0.93	−0.82***	0.41**	3.41
Term spread (TMS)	−0.41**	0.89	−0.81***	−0.39**	3.31
Default yield spread (DFY)	−0.16	0.13	−0.84***	−0.21	2.74
Default return spread (DFR)	0.36	0.68	−0.80***	0.32	3.06
Inflation rate (INFL)	0.01	0.00	−0.83***	0.05	2.54
Output gap (OG)	−0.46***	1.08	−0.75***	−0.33**	3.07
Short interest (SI)	−0.55**	1.48	−0.84***	−0.46*	3.85
Kitchen sink (via elastic net)	–	–	−0.72***	–	5.50

Table 5
Controlling for uncertainty measures.

Panel A presents the correlations between the disagreement index and eight uncertainty measures, including economic uncertainty (Bali, Brown, and Caglayan, 2014), treasury implied volatility (Choi, Mueller, and Vedolin, 2017), financial uncertainty and macro uncertainty (Jurado, Ludvigson, and Ng, 2015), economic policy uncertainty (Baker, Bloom, and Davis, 2016), news implied volatility (Manela and Moreira, 2017), sample variance (Welch and Goyal, 2008), and the Chicago Board Options Exchange (CBOE) volatility index (VIX). Panel B presents the results of predicting market returns with one uncertainty measure as

$$R_{t+1} = \alpha + \psi U_t + \varepsilon_{t+1}.$$

Panels C presents the results of predicting market returns with the disagreement index and one uncertainty measure as

$$R_{t+1} = \alpha + \beta D_t + \psi U_t + \varepsilon_{t+1}.$$

The last row reports the slope of the disagreement index from an elastic net regression by including all the uncertainty measures, where the t-value is calculated following Tibshirani, Taylor, Lockhart, and Tibshirani (2016) and Lee, Sun, Sun, and Taylor (2016). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	Correlation	Univariate		Bivariate		
Uncertainty	Corr(Uncertainty$_t$, D_t)	ψ	R^2	β	ψ	R^2
Economic uncertainty	0.09	−0.13	0.09	−1.02***	−0.04	4.88
Treasury implied volatility	0.33***	−0.37	0.70	−1.01***	−0.06	4.72
Financial uncertainty	0.23***	−0.62**	2.01	−0.69***	−0.49*	3.68
Macro uncertainty	0.24***	−0.45	1.06	−0.74***	−0.30	2.96
Economic policy uncertainty	−0.20***	0.25	0.32	−0.86***	0.10	3.02
News implied volatility	0.07	0.09	0.04	−0.85***	0.14	2.74
Sample variance	0.23***	−0.22	0.26	−0.81***	−0.07	2.54
VIX	0.26***	0.00	0.00	−1.06***	0.24	4.66
Kitchen sink (via elastic net)	–	–	–	−0.89**	–	5.09

Table 6
Asset allocation results.

This table reports portfolio gains of a mean-variance investor with risk-aversion $\gamma = 3$ or 5 for predicting market returns with the PCA, equal-weight (individual measure), and PLS disagreement indexes, respectively. The investor allocates her wealth monthly between the stock market and the risk-free asset by applying the out-of-sample forecasts based on one of the three disagreement indexes. CER gain is the annualized certainty equivalent return difference between applying a disagreement index forecast and applying the historical return mean forecast. Sharpe ratio is the monthly average portfolio excess return divided by its standard deviation. The portfolio weight is estimated recursively using data available at the forecast formation month t. The investment period is 1991:02–2018:12. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	No transaction cost		50 bps transaction costs	
	CER gain (%)	Sharpe ratio	CER gain (%)	Sharpe ratio
Panel A: Risk aversion $\gamma = 3$				
PCA	0.71	0.14**	0.52	0.13
Equal-weight	1.70**	0.16**	1.33*	0.15**
PLS	2.50***	0.18***	1.92**	0.16**
Panel B: Risk aversion $\gamma = 5$				
PCA	0.96**	0.12**	0.81**	0.11**
Equal-weight	2.10***	0.16***	1.69**	0.14**
PLS	2.68***	0.17***	1.88**	0.14**

Table 7
Out-of-sample R^2_{OS}s of forecasting market returns with alternative methods.

This table presents the out-of-sample R^2_{OS}s of forecasting h-month-ahead market returns with six alternative information aggregation methods: equal-weight LASSO, combination LASSO, encompassing LASSO, adaptive LASSO, egalitarian LASSO, and elastic net. Statistical significance for R^2_{OS} is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Method	$h = 1$	$h = 3$	$h = 12$
Equal-weight LASSO	1.26^{**}	6.09^{***}	12.08^{***}
Combination LASSO	1.08^{**}	2.87^{***}	9.67^{***}
Encompassing LASSO	1.09^{**}	2.92^{***}	9.41^{***}
Adaptive LASSO	0.71^{*}	2.34^{***}	7.47^{***}
Egalitarian LASSO	1.30^{*}	2.69^{***}	8.42^{***}
Elastic net	1.36^{**}	2.69^{***}	8.43^{***}

Table 8
Asymmetric forecasting power of disagreement.

Panel A reports the in- and out-of-sample R^2s of predicting market returns with the disagreement index in different time periods, which are calculated as in Eq. (19). Panel B presents the results of predicting market returns with a state-dependent regression:

$$R_{t+1} = \alpha + \beta_{\text{high}} S_t^{\text{high}} D_t + \beta_{\text{low}} S_t^{\text{low}} D_t + \varepsilon_{t+1},$$

where S_t^{high} (S_t^{low}) is a dummy variable that equals one if month t is in high- (low-) sentiment periods and zero if month t is in low- (high-) sentiment periods (Baker and Wurgler, 2006). Statistical significance for R^2_{OS} is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing $H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Forecasting performance in different periods

In-sample R^2		Out-of-sample R^2_{OS}	
High sentiment	Low sentiment	High sentiment	Low sentiment
5.28	0.80	3.69**	−0.55

Panel B: State-dependent regression

	β_{high}	t-value	β_{low}	t-value	R^2
Sentiment-based state	−1.12***	−4.71	−0.42	−1.21	2.96

Table 9
Disagreement and expectations of market returns.

Panel A reports the results of regressing expectations of market returns on the disagreement index:

$$\text{Expectation}_{t:t+12} = \alpha + \beta D_t + \varepsilon_t,$$

where $\text{Expectation}_{t:t+12}$ is investor expectation of 12-month-ahead market return at time t, which is measured by (value-weighted) aggregate analysts' return forecast (Engelberg, McLean, and Pontiff, 2020), the University of Michigan Surveys of Consumers Attitudes, Graham-Harvey's survey of CFOs, or Robert Shiller's survey of individual investor confidence. Panel B reports the results of regressing return forecast errors on the PLS disagreement index:

$$\text{Realized return}_{t:t+12} - \text{Expectation}_{t:t+12} = \alpha + \beta D_t + \varepsilon_t,$$

where the $\text{Expectation}_{t:t+12}$ in the case of Shiller's survey is the projection of analysts' return forecast on Shiller's survey, so that the projection has the same measurement unit as the realized return. The sample periods all end in 2018:12, but start differently, from 1999:04 for analysts' return forecast, 2002:06 for the University of Michigan's survey, 2000:10 for Graham-Harvey's survey, and 2001:07 for Shiller's survey, respectively. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Expectations of market returns				
	$\text{Corr}(\text{Expectation}_{t:t+12}, D_t)$	β	t-value	R^2
Analysts' return forecast	0.35***	3.26***	2.61	12.45
University of Michigan survey	0.24***	2.71*	1.69	5.55
Graham-Harvey's survey	0.26**	0.57**	2.40	6.58
Shiller's survey	0.25***	2.16**	2.31	6.50
Panel B: Market return forecast errors				
		β	t-value	R^2
Analysts' return forecast		−7.34***	−2.73	23.26
University of Michigan survey		−8.42***	−3.36	21.17
Graham-Harvey's survey		−9.65***	−4.41	31.61
Shiller's survey		−9.18***	−4.45	30.34

Table 10
Forecasting market returns with cash flow news- and discount rate news-based disagreement indexes.

This table reports the slopes and Newey-West t-values from the regression of

$$R_{t+1} = \alpha + \beta_{\mathrm{CF}} D_t^{\mathrm{CF}} + \beta_{\mathrm{DR}} D_t^{\mathrm{DR}} + \varepsilon_{t+1},$$

where D_t^{CF} (D_t^{DR}) is the PLS disagreement index that uses the contemporaneous cash flow news (discount rate news) as the regressor in Eq. (10). The cash flow news and discount rate news are estimated by using the Campbell (1991) VAR approach. In the leftmost column, "$\tilde{R}$" represents the total market return, economic variables are defined in Table 4, and "PC" represents the first three principal components extracted from all the economic variables (except for DP). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

VAR variables	β_{CF}	t-value	β_{DR}	t-value	R^2
$\tilde{R}$, DP	-0.65^{***}	-2.92	-0.01	-0.08	2.15
$\tilde{R}$, DP, DY	-0.65^{***}	-3.08	-0.01	-0.03	2.16
$\tilde{R}$, DP, EP	-0.67^{***}	-2.73	-0.11	-0.51	2.06
$\tilde{R}$, DP, DE	-0.67^{***}	-2.73	-0.11	-0.51	2.06
$\tilde{R}$, DP, RVOL	-0.70^{**}	-2.52	-0.06	-0.23	2.23
$\tilde{R}$, DP, BM	-0.65^{***}	-3.29	0.11	0.61	2.16
$\tilde{R}$, DP, NTIS	-0.66^{***}	-3.02	-0.02	-0.11	2.21
$\tilde{R}$, DP, TBL	-0.55^{***}	-2.86	0.10	0.59	1.62
$\tilde{R}$, DP, LTY	-0.64^{***}	-3.23	0.09	0.53	2.13
$\tilde{R}$, DP, LTR	-0.65^{***}	-3.13	0.02	0.09	2.24
$\tilde{R}$, DP, TMS	-0.57^{***}	-2.87	0.07	0.39	1.81
$\tilde{R}$, DP, DFY	-0.65^{***}	-2.82	-0.02	-0.11	2.16
$\tilde{R}$, DP, DFR	-0.65^{***}	-2.70	-0.03	-0.13	2.10
$\tilde{R}$, DP, INFL	-0.65^{***}	-3.29	0.03	0.16	2.20
$\tilde{R}$, DP, OG	-0.57^{***}	-2.83	-0.06	-0.32	1.57
$\tilde{R}$, DP, SI	-0.47^{**}	-2.08	0.27	1.28	1.93
$\tilde{R}$, DP, PC	-0.66^{***}	-2.78	0.04	0.16	2.26

Table 11
Relation of disagreement with market volatility and trading volume.

Panel A presents the results of predicting the volume-volatility elasticity with the disagreement index:

$$\text{Elasticity}_{t+1} = \alpha + \beta D_t + \varepsilon_{t+1},$$

where the elasticity in month $t+1$ is the slope of regressing the daily change in turnover of NYSE stocks on the daily change in volatility within month $t+1$. Realized volatility, realized semi-volatility, and median realized volatility are estimated based on the S&P 500 index returns from five-minute intervals (Andersen, Dobrev, and Schaumburg, 2012), and futures realized volatility is estimated based on the S&P 500 index futures contract returns from five-minute intervals (Johnson, 2019). Panel B presents the results of the following regression:

$$\text{Volume}_{t+1} = \alpha + \beta_1 \text{D\_Volatility}_t + \beta_2 \text{Volatility}_t^{\circ} + \varepsilon_{t+1}.$$

D_Volatility is the disagreement-related volatility and is extracted using the PLS method, and Volatility$^{\circ}$ is the residual of regressing volatility on D_Volatility. Following Hamilton (2018), we apply AR(4) to both trading volume and market volatility to remove potential trends and expected information. Reported are the regression coefficient, Newey-West t-value, and R^2. The sample period is 2000:01–2018:12 for the first three volatility measures and 1990:01–2015:12 for the last one. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Predicting volatility-volume elasticity			
Volatility measure	β	t-value	R^2
Realized volatility	4.01***	4.41	4.94
Realized semi volatility	1.85**	2.08	1.83
Median realized volatility	1.99**	1.98	1.25
Futures realized volatility	2.24**	2.17	1.49

Panel B: Predicting trading volume					
Volatility measure	β_1	t-value	β_2	t-value	R^2
Realized volatility	2.38***	3.32	−1.30	−1.36	5.31
Realized semi volatility	2.53***	3.44	−1.16	−1.32	5.58
Median realized volatility	2.60***	3.31	−1.44*	−1.77	6.36
Futures realized volatility	1.13**	2.15	−0.46	−0.75	1.18