


Article

“Agree to Disagree”: Forecasting Stock Market Implied Volatility Using Financial Report Tone Disagreement Analysis

Nicolas S. Magner ^{1,*}, Nicolás Hardy ¹, Tiago Ferreira ² and Jaime F. Lavin ³ ¹ Facultad de Administración y Economía, Universidad Diego Portales, Santiago 8370191, Chile² Facultad de Economía y Negocios, Universidad Alberto Hurtado, Santiago 6500620, Chile³ Escuela de Negocios, Universidad Adolfo Ibáñez, Santiago 7941169, Chile

* Correspondence: nicolas.magner@udp.cl

Abstract: This paper studies the predictability of implied volatility indices of stocks using financial reports tone disagreement from U.S. firms. For this purpose, we build a novel measure of tone disagreement based on financial report tone synchronization of U.S. corporations scattered across five Fama-French industries. The research uses tree network methods to calculate the minimum spanning tree length utilizing data from text mining sentiments features extracted from all U.S. firms that considers 837,342 financial reports. The results show that periods of increased disagreement predict higher implied volatility indices. We contribute to the literature that proposes that a high level of expectations dispersion leads to higher stock volatility and fills a gap in understanding how firms' disagreement level of financial report tone forecast the aggregate stock market behavior. The findings also have implications for financial stability and delegated portfolio management, as accurate volatility prediction is critical for practitioners.

Keywords: disagreement; textual analysis; predictability; stock returns; implied volatility; network methods; forecast models

MSC: 91G45; 91G70; 91G15



Citation: Magner, N.S.; Hardy, N.; Ferreira, T.; Lavin, J.F. “Agree to Disagree”: Forecasting Stock Market Implied Volatility Using Financial Report Tone Disagreement Analysis. *Mathematics* **2023**, *11*, 1591. <https://doi.org/10.3390/math11071591>

Academic Editors: Christos Floros, Christos Kountzakis and Konstantinos Gkillas

Received: 30 January 2023

Revised: 18 February 2023

Accepted: 20 February 2023

Published: 25 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As a consequence of the disagreement regarding the fundamentals of companies and the economy among investors, the stock market significantly captures the attention of practitioners and academics. However, the rational behavior assumptions that underlie the main asset pricing models show gaps since rational behavior only sometimes becomes observed at the level of individual agents' behavior and the aggregate market level [1]. To properly tackle the above, the behavioral finance literature developed approaches of representative-agent models with rational expectations and unconventional preferences, representative-agent models with standard preferences but biased beliefs, and a variety of heterogeneous-agent models [2–4].

Disagreement models, a particular class of heterogeneous-agent models, address different mechanisms to explain conflicting behaviors with the assumption of rationality prevailing among classical asset pricing models and prices and volume patterns observed in the financial markets (see [5,6] for surveys). For instance, mechanisms related to gradual information flow, limited attention, and heterogeneous priors explain specific patterns of predictability in stock markets associated with returns, volatility, and trading volume. These models, at empirical and theoretical levels, exemplify the importance of differences in financial agents' expectations as market behavior factors [1].

The analysis of the role played by investor disagreement on market price can help to understand how investors update their expectations. Some standard theoretical approaches assume investors agree on the interpretation of signals and use prices to update their expectations while in disagreement models investors can “agree to disagree” in a different

opinion to update their expectations [7]. Despite a substantial body of research examining the implications of financial report tone on the market, there is still a gap in the literature examining the effect of disagreement level on stock market behavior. Specifically concerning the relationship between tone disagreement and the market's aggregate implied volatility.

Previous research suggests that positive language is associated with a firm's current performance and lower return volatility [8,9] and also with the firm's future performance [8,10–12] while a negative sentiment in annual reports is linked to larger delisting probabilities, lower odds of paying dividends, higher loan loss provisions, and lower future return on assets [13]. Notwithstanding previous results finding evidence in specific industries and firms, our research sheds light on the effects of disagreement on the stock market in terms of its future implied volatility. The relationship between firms' disagreement level of financial report tone on the aggregate stock market implied volatility is precisely the gap that this paper aims to fill.

In this study, we find that disagreement in tone regarding firms' perspectives within financial reports from U.S. companies can predict implied volatility of stock indices. Additionally, we find that periods of increased disagreement predict higher implied volatility indices. These results are in line with literature that analyzes how expectations dispersion among market agents influences the behavior of stock markets. Precisely, our findings contribute to empirical and theoretical literature that proposes that a high level of expectations dispersion leads to higher stock volatility [14–20]. Therefore, we argue that when the tone varies widely between firms, investors collectively disagree more about the future market outcome, resulting in greater implied volatility in the stock market.

In this paper, we use the Minimum Spanning Trees (MST), a correlation network methodology used to study the increase in cointegration between financial markets. With this method, we calculated the Minimum Spanning Tree Length (MSTL) to measure the synchronization of disagreement in tone regarding firms' perspectives within financial reports from U.S. companies. This measure is a parsimonious representation of the complex network of interrelationships, and the help of its connections obtain direct and indirect information about the firm-specific tone. Additionally, this methodology represents a dynamic system for measuring the complex relationship between market information and asset pricing [21].

Our findings are important for a couple of reasons. First, we present a new method for measuring disagreement by analyzing the direct disclosure of all 10-K/Q financial reports from all U.S. firms and applying a network approach to estimate tone synchronization. To achieve this, we measured the monthly financial report correlation level to build the Minimum Spanning Tree Length (MSTL) [22] that allows us to measure the tone synchronization of all firms classified in five different Fama-French industries [23]. Secondly, we provide evidence of the relevance of information presented in financial reports. It offers insight into the firm's perspective to forecast future stock market behavior via implied volatility forecasting. Being the latter an essential activity among financial agents, given that accurate volatility prediction is a critical task for many practitioners involved in financial stability and delegated portfolio management.

We contribute to several aspects of the literature on disagreement and stock market behavior. Specifically, using a network approach, our study takes soft information in financial reports to quantify the degree of disagreement among investors and predict implied volatility indices. This approach has several benefits when compared to using analyst forecast dispersion. First, financial reports are publicly available for all listed firms on the SEC's EDGAR website, whereas analyst forecasts may only be available for a subset of companies. This issue allows for a more comprehensive analysis of disagreement among investors. Secondly, only some companies have analyst coverage, as analysts tend to cover larger firms. Using financial reports allows for a broader measure of disagreement that includes firms of different sizes and types and therefore captures a more diverse group of investors. Thirdly, it allows for a broader measure of disagreement that includes distinct

types of firms and investors. This approach captures a more diverse group of investors and provides a more comprehensive understanding of investor sentiment.

Our empirical strategy also contributes to articulating methods to study the phenomenon of disagreement. Previous findings investigating the impact of financial report tone on market volatility and return are intriguing. However, it is impossible to do real-time forecasting with in-sample analysis with panel models, especially given the scrutiny that in-sample exercises are prone to data mining-induced overfitting [24]. To fill this void, in this paper, we focus on out-of-sample analyses with various loss functions, such as mean squared prediction error [25] and mean directional accuracy [26]. Accordingly, we demonstrate that the current month's tone disagreement level provides important information regarding the month ahead return of volatility indices.

In addition to out-of-sample analysis, we employed mean directional accuracy to develop a robustness test. With this strategy, we assess the MSTL tone's ability to anticipate the direction of change in implied volatility indices. Thus, independent of the forecast error, we examine the accuracy of our forecasts in estimating whether implied volatility indices would rise or decrease in the forthcoming month [27,28].

Finally, we contribute to enhancing volatility index predictions using public financial report data and to the literature on disseminating public financial report information [29–31]. We employ a novel specification that considers the date that financial reports are provided to the SEC as part of a conditional expectation set of information that exploits aggregate information accessible to investors at time t to estimate the $t + 1$ implied volatility indices variations.

To the best of our knowledge, this is the first study to examine the relations between future implied volatility index returns and aggregated firm tone disagreement from financial reports. Our research suggests a new potential route for establishing a time series link between overall non-numerical financial report data and volatility indices returns. Furthermore, we propose an innovative way of applying network analysis methodologies to measure “corporate financial disagreement.” However, it is essential to mention that this paper does not investigate the structural or channel relationships between “corporate financial report tone disagreement” and the prices of volatility indices. In this spirit, we make no empirical “causation” claims. According to Rossi [32], such an analysis would necessitate the use of a structural model, which is beyond the scope of this paper.

The remainder of the paper is structured as follows: Section 2 reviews the literature, Section 3 describes the data and forecasting models used, Section 4 presents in-sample and out-of-sample results, Section 5 discusses the results and implications, and Section 6 summarizes our conclusions.

2. Literature Review

The role of investor disagreement explains how investors update their expectations, impacting trading volumes and prices, particularly returns and volatility. Standard theoretical approaches assume rational expectations, where investors agree on the interpretation of signals and use prices to update their expectations. On the contrary, in disagreement models, investors do not use prices but “agree to disagree” in a different opinion to update their expectations [7]. These differences of opinion between investors generate different patterns in the prices of financial assets. Studies evidence that higher disagreement among investors leads to mixed results on returns but to a higher volume and higher volatility [14]. Similarly, Ref. [14] contends that when investors have infrequent but substantial disagreements, there is a positive autocorrelation in volume and a positive correlation between volume and volatility.

There is firm-specific and external information to study the role of investor disagreement. However, the literature that analyzed tone disagreement regarding the firm's future and market performance mainly used external sources of information rather than financial reports. For example, Dzieliński and Hasseltoft [33] measure aggregate disagreement through linguistic analysis of firm-specific news and found that news tone dispersion is

positively related to volatility and turnover. Cookson and Niessner [34] study the sources of investor disagreement using the sentiment of investors from StockTwits, a social media investing platform, suggesting that information differences are even more important for trading than differences across market approaches.

Concerning the role of investor disagreement, there still needs to be more research considering financial reports as a source of disagreement. Financial reports can be a potential source of disagreement for investors since this is where the firm states many opinions regarding future firm performance. Financial reports for companies with greater present and future performance and lower return volatility, according to Li [21], tend to use more positive tone language. Gandhi et al. [13] show that negative sentiment in annual reports of U.S. banks is linked to larger delisting probabilities, lower odds of paying dividends, higher loan loss provisions, and lower future return on assets. Furthermore, short-term market reactions to Forms 10-Q and 10-K filings are more positive when the tone of these disclosures is more favorable [35].

Huang et al. [12] show that managers can use tone to mislead investors about firm fundamentals, documenting that an abnormal level of positive tone by managers predicts negative future earnings and cash flows. Davis and Tama-Sweet [8,10–12] found that firms disclose more optimistic language in earnings press releases than in the MD&A section of 10-K reports, implying that managers use tone to manage investor expectations by omitting or shifting the pessimistic language from earnings press releases when they face incentives to report strategically.

There is ample evidence relating differences in tone and market participants' reactions. For instance, studies on the wording of CEO letters to the public have found a connection between greater confidence and higher stock returns [36,37], and greater negativity in these letters is associated with lower stock returns [38]. Kothari et al. [39] show that positive disclosures decrease the cost of capital and return volatility. Based on the evidence above, it is straightforward to conclude that managers may be influenced by incentives when deciding the tone they will use in their financial reports, which is understandable given that tone has been shown to influence investor behavior.

Overall, the literature mentioned above has documented significant impacts of financial report tone on volatility, which is a very relevant finding given that stock market volatility has a significant impact on the efficiency of asset pricing, portfolio, and derivative hedging strategies, as well as financial risk management [40]. However, it is challenging to increase prediction accuracy, making volatility forecasting a complex and important research topic [40]. Implied stock volatility indexes, in particular, serve a vital function as the investor fear gauge, an important indicator of market sentiment [41].

Banerjee et al. [15] find a negative relationship between the S&P 500 performance and the VIX evolution. Similarly, Antonakakis et al. [42] document a negative correlation between U.S. stocks' performance and the VIX's evolution. Additional evidence shows that implied volatility indices fluctuations impact markets return in local terms but also have a leading role in international stock markets [43,44]. However, given what has been discussed thus far, there still needs to be a gap in understanding the relationship between firms' disagreement level of financial report tone and the market's aggregated implied volatility.

3. Materials and Methods

3.1. Implied Volatility Indices Database

We consider monthly closing prices of four implied volatility indices (CBOE Russell 2000 Volatility Index (RVX), CBOE Volatility Index (VIX), Cboe DJIA Volatility Index SM (VXD), and CBOE S&P 100 Volatility Index (VXO)). Our sample goes in monthly frequency from January 2000 to December 2022. We download our series from Refinitiv Thomson Reuters and Bloomberg financial services platforms.

3.2. Financial Report Tone

We use the financial report sentiment word counts database from Bill McDonald (<https://sraf.nd.edu/>, accessed on 1 October 2022) to measure the tone of each 10-K/Q. Specifically, we include the following reports in our sample: 10-K, 10-K-A, 10-K405, 10-K405-A, 10-KT, 10-KT-A, 10-Q, 10-Q-A, 10-QT, 10-QT-A, 10KSB, 10KSB-A, 10KSB40, 10KSB40-A, 10QSB and 10QSB-A. Finally, we base our analysis from 2001 to 2021, with 837,342 reports considered.

Financial report tone is the difference between the net positive word count and the negative word count, scaled by the count of all words multiplied by 10. Positive and negative words are defined based on Tim Loughran and Bill McDonald seminal paper [45].

The net positive word is the number of positive words minus the number of negation words, counted when they occur within four or fewer words of a positive word. Negation terms include “no, not, none, neither, never, and nobody”. For a complete list of positive and negative words, please refer to <https://sraf.nd.edu/loughranmcdonald-master-dictionary/> (accessed on 1 October 2022).

We estimate an average daily tone using all financial reports for each of the five Fama-French industries. In the next section, we present our network analysis methodology based on estimating the Minimum Spanning Tree Length (MSTL) over each industry’s five Fama-French industries’ daily tone average.

3.3. Financial Report Tone Disagreement: A Network Analysis Approach

We use network analysis method to measure financial report tone disagreement. In specific, we employed the minimum spanning tree length (MSTL) methodology to represent the correlation phenomenon between the tone of the five economic sectors considered.

Following [46–51] we used the correlations between the sector’s tones and build the asset trees. For this purpose, we take the daily tone average of the sector i extracted from the financial reports of the firms belonging to it, and calculate monthly correlations for each pair of the sectors i and j (See Equation (1)).

Let \mathbf{tone}_i^t be the sectors tone vector of the day t' , then:

$$\rho_{ij}^t = \frac{\langle \mathbf{tone}_i^{t'} \mathbf{tone}_j^{t'} \rangle - \langle \mathbf{tone}_i^{t'} \rangle \langle \mathbf{tone}_j^{t'} \rangle}{[\langle \mathbf{tone}_i^{t'2} \rangle - \langle \mathbf{tone}_i^{t'} \rangle^2][\langle \mathbf{tone}_j^{t'2} \rangle - \langle \mathbf{tone}_j^{t'} \rangle^2]} \tag{1}$$

ρ_{ij}^t is the Pearson’s correlation coefficient between the sector tone i and j where $\langle \cdot \rangle$ indicates the average over all the trading days of the month t . In this way, a $N \times N$ symmetrical matrix \mathbf{C}^t of correlations between sector’s tone (N is the number of sectors) with values $-1 \leq \rho_{ij} \leq 1$.

Then, we convert the correlations of \mathbf{C}^t in distances $d_{ij} = (2(1 - \rho_{ij}))^{1/2}$ to represent the distance between the sectors i and j . Thus, a correlation $\rho_{ij} = -1$ indicates a maximum distance of $d_{ij} = 2$, while $\rho_{ij} = 1$ indicates a minimum distance of $d_{ij} = 0$.

With the complete correlation network, we used the Prim algorithm [52] to estimate the Minimum Spanning Tree (hereafter, MST) to reduce the information space of the entire network by connecting all nodes with $N(N - 1)/2$ edges, to a tree with $N - 1$ edges. Finally, we estimate the normalized length of the MST ($MSTL$) as:

$$MSTL_t = \frac{1}{N - 1} \sum_{d_{ij} \in \mathbf{T}^t} d_{ij}^t \tag{2}$$

As the sum of the edges of the resulting tree \mathbf{T}^t calculated for each month t forms a time series. The variation in the $MSTL$ is calculated as $\Delta MSTL_t = \ln MSTL_t - \ln MSTL_{t-1}$, which allows us to work with a stationary time series. With this specification, we obtain a dynamic tone disagreement measure tree based on their correlations.

3.4. Forecasting Methods

To evaluate the forecasting power of our disagreement variable $[\Delta MSTL_t]$ we employed two econometrics methodologies. First, we design an in-sample analysis using econometrics models displayed on Table 1 Row 1. Second, to avoid overfitting, we implement an out-of-sample tests comparing the out-of-sample core model (See Table 1, row 1) with the benchmark model (Table 1, row 2). Following [32,36,37] we build our benchmark model considering autoregressive components of implied volatility index since $AR_{(p)}$ models are difficult benchmarks to outperform.

Table 1. The main econometric models for both the core and benchmark models.

In-the-Sample and Out-of-Sample Core Model	
$\Delta \ln(P_{i,t}) = c + \beta \Delta MSTL_{t-1} + \gamma_i \sum_{n=1}^3 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$	(1)
Out-of-sample benchmark models	
$\Delta \ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^3 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$	(2)
$\Delta \ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^6 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$	(3)

Source: authors' elaboration.

In Table 1, $\Delta MSTL_t$ is the log-difference of financial reports tone synchronization at time t . $P_{i,t}$ is the implied volatility index i at time t , where n is the number of lags (either 3 or 6). Finally, $\varepsilon_{i,t}$ are error terms. To assess the predictive ability of the $\Delta MSTL$, we evaluate the null hypothesis $H_0 : \beta = 0$; a rejection of the null hypothesis means that the $\Delta MSTL$ is providing valuable information to predict the implied volatility indices $P_{i,t}$. To conduct our in-sample analyses, we test H_0 through a simple t statistic, with a HAC estimator of the long run variance, following Newey and West [53,54].

Additionally, we included multiple unknown break analyses to account instabilities in the predictive performance of forecasting models. Timmermann [55] found evidence of sporadic and unstable predictability, appearing as “pockets of predictability”. For this reason, we allow the possibility of multiple unknown breaks in the parameters in the core in-sample model (See Table 1, Row 1). Simply stated, this is a simple case of time-varying coefficients model [56]. We allowed a maximum of five breaks with maximum of six regimes. We determine each regime breakpoints using the UDmax test of [57].

3.4.1. Encompassing Test ENCNEW

For out-of-sample evaluations with nested models we use the Clark and McCracken ENCNEW test [58] with a recursive scheme to update our parameters. Under the null hypothesis of no predictability, the asymptotic distribution of the ENCNEW is non-standard, and depends on how parameters are updated (rolling or recursive windows), the number of excess parameters in the large nesting model (in our case only one, the first lag of the tone MSTL), and the P/R ratio (where P is the number of observations in the prediction window), and R the number of observations in the estimation window [58,59]:

$$ENCNEW = P \frac{P^{-1} \sum_{t=R}^{T-1} \hat{e}_{1,t+1} (\hat{e}_{1,t+1} - \hat{e}_{2,t+1})}{P^{-1} \sum_{t=R}^{T-1} \hat{e}_{2,t+1}^2} \tag{3}$$

where $\hat{e}_{1,t+1}$ is the forecasting error of the benchmark model, and $\hat{e}_{2,t+1}$ is the forecasting error of the proposed nesting model.

Clark and McCracken [58] simulate the asymptotic distribution of the ENCNEW and establish the relevant critical values for different number of excess parameters (in our case $k_2 = 1$), different schemes to estimate our parameters (in our case a recursive scheme), and different divisions of the data base (in our case $P/R = 0.4, 1, 2,$ and 4). The critical values can be found in the notes of the tables.

3.4.2. Encompassing T-Test (ENC-t)

Another traditional out-of-test is the ENC-t, proposed by Clark and McCracken [58]. Akin to the ENCNEW, the asymptotic distribution of the ENC-t, under the null hypothesis, is non-standard, and depends on the same parameters. In particular, the ENC-t is just:

$$ENC - t = (P - 1)^{0.5} \frac{P^{-1} \sum_{t=R}^{T-1} \hat{e}_{1,t+1} (\hat{e}_{1,t+1} - \hat{e}_{2,t+1})}{\sqrt{\hat{S}}} \tag{4}$$

where \hat{S} is simply the standard variance of $\hat{e}_{1,t+1} (\hat{e}_{1,t+1} - \hat{e}_{2,t+1})$ (not the HAC estimator).

The criticals values of the test depend on the number of excess parameters in the nesting model (one in our case), on the scheme used for updating our parameters (recursive windows in our case), and on the ratio P/R ratio (4, 2, 1, and 0.4 in our case).

3.4.3. Mean Directional Accuracy

Following to [27,28] we assess the ability of the MSTL tone in anticipating the direction of change of implied volatility indices. In this test we examine the accuracy of our forecasts in estimating whether implied volatility indices would rise or decrease in the forthcoming month:

$$W_t = \begin{cases} 1 & \text{if } (\Delta \ln(P_t))(f_{t-1}) > 0 \\ 0 & \text{if } (\Delta \ln(P_t))(f_{t-1}) \leq 0 \end{cases} \tag{5}$$

where f_{t-1} represents a generic forecast for the one-period change of implied volatility index $\Delta \ln(P_t)$. Our variable computes a “hit” whenever our forecast predicts correctly de direction of $\Delta \ln(P_t)$. We are interested in evaluating the null hypothesis $H_0 : E(W_t) \leq 0.5$; in other words, this is equivalent to testing wheteher our forecast correctly predicts the sign of $\Delta \ln(P_t)$ more times than a simple “fair coin”. We evaluate this null hypothesis through the Diebold and Mariano [60] test. Results using the Pesaran and Timmermann [61] are qualitatively the same, and they are available upon request.

In an additional analysis, we explore the ability of forecast predicting the sign of $\Delta \ln(P_t)$ compared to the nested benchmark (see Table 1, row 1). To this end, we define the following dummy variable Z_t :

$$Z_t = \begin{cases} 1 & \text{if } (\Delta \ln(P_t))(g_{t-1}) > 0 \\ 0 & \text{if } (\Delta \ln(P_t))(g_{t-1}) \leq 0 \end{cases} \tag{6}$$

where g_t is simply the generic forecast of the benchmark nested model. Then we evaluate the null hypothesis $H_0 : E(W_t) \leq E(Z_t)$ through a simple Diebold and Mariano [60] test. A rejection of the null hypothesis indicates that our core model (See Table 1, row 2) statistically outperform the benchmark when predicting the sign of the implied volatility index.

4. Results

In this section, we report and discuss our in-sample results, considering the entire window period (Table 2) and regime breaks (Table 3), and the following out-of-sample test results: ENCNEW test (Table 4), ENC-t test (Table 5), and Mean Directional Accuracy test (Table 6).

Table 2. Forecasting stock market implied volatility with financial reports tone disagreement: In-sample analysis.

	(1) VIX	(2) RVX	(3) VXD	(4) VXO
$\Delta MSTL_{t-1}$	0.197 ** (0.093)	0.154 * (0.092)	0.186 ** (0.086)	0.227 ** (0.104)
AR(1)	−0.225 ** (0.094)	−0.163 (0.117)	−0.203 ** (0.097)	−0.155 * (0.084)
AR(2)	−0.110 ** (0.056)	−0.064 (0.054)	−0.085 (0.053)	−0.122 ** (0.055)
AR(2)	−0.131 *** (0.040)	−0.160 *** (0.047)	−0.134 *** (0.039)	−0.125 *** (0.039)
Constant	−0.001 (0.008)	0.001 (0.008)	−0.001 (0.008)	−0.002 (0.009)
R-squared	0.076	0.058	0.068	0.061
N	261	212	261	261

Notes: Each panel report estimates of parameters based on Equation (1). * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$. Source: Authors elaboration.

Table 3. Forecasting stock market implied volatility with financial reports tone disagreement: In-sample analysis with regime breaks.

	(1) VIX	(2) RVX	(3) VXD	(4) VXO
Regime 1				
$\Delta MSTL_{t-1}$	0.204	0.324 ***	0.015	0.151
Time period	2000M01–2003M06	2004M05–2008M09	2000M01–2003M06	2000M01–2003M06
Regime 2				
$\Delta MSTL_{t-1}$	0.311 ***	0.009	0.379 **	0.292 *
Time period	2003M07–2007M04	2008M10–2012M05	2003M07–2007M04	2003M07–2007M04
Regime 3				
$\Delta MSTL_{t-1}$	0.351 *	0.018	0.304 *	0.424 **
Time period	2007M05–2010M07	2012M06–2015M08	2007M05–2010M07	2007M05–2010M07
Regime 4				
$\Delta MSTL_{t-1}$	−0.200 *	0.292 **	−0.218 **	−0.2051 *
Time period	2010M08–2013M12	2015M09–2018M11	2010M08–2013M12	2010M08–2013M12
Regime 5				
$\Delta MSTL_{t-1}$	0.177	0.192	0.360 ***	0.205
Time period	2014M01–2017M07	2018M12–2021M12	2014M01–2017M03	2014M01–2017M10
Regime 6				
$\Delta MSTL_{t-1}$	0.433 *		0.304	0.630 ***
Time period	2017M08–2021M12		2017M04–2021M12	2017M11–2021M09
R ²	0.2	0.226476	0.211	0.183068
N	261	212	261	261

Notes: Breaks are determined according to UD-Max statistic by [57]. Each panel report estimates of our parameters for a different regime based on Equation (1). We do not report the autoregressive components, the constant nor the standard deviations to save space. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$. Source: Authors elaboration.

Table 4. Forecasting stock market implied volatility with financial reports tone disagreement: Out-of-sample analysis with the ENCNEW test.

Contrasting against an AR(1) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \Delta \ln(P_{i,t-1}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	0.952	4.122 **	2.253 *	4.596 **
2	−0.439	2.939 **	1.664 *	3.461 **
1	0.668	1.656 **	0.434	2.066 **
0.4	0.679	1.379 **	1.024 *	1.867 **
Contrasting against an AR(2) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^2 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	0.754	3.503 **	1.936 *	3.752 **
2	−0.527	2.506 **	1.507 *	2.882 **
1	0.580	1.363 *	0.272	1.643 **
0.4	0.611	1.193 **	0.915 *	1.559 **
Contrasting against an AR(3) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^3 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	0.686	3.351 **	1.679 *	3.589 **
2	−0.477	2.411 **	1.306 *	2.773 **
1	0.649	1.348 *	0.304	1.618 **
0.4	0.652	1.080 **	0.817 *	1.442 **
Contrasting against an AR(6) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^6 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	1.236	3.387 **	2.176 *	3.110 **
2	0.028	2.511 **	1.631 *	2.573 **
1	0.877	1.301 *	0.541	1.382 *
0.4	0.925 *	0.992 *	0.964 *	1.248 **

Notes: Each entry reports the EncNew statistic with a recursive scheme. This table compares the predictive ability of our out-of-sample core models (See Table 1, row 2) with the predictive ability of the benchmark models (See Table 1, row 3). We report four estimation windows with the following critical values: P/R = 4 used critical values 1.540, 2.561, and 5.087 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 2 used critical values 1.280, 2.085, and 4.134 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 1 used critical values 0.984, 1.584, and 3.209 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 0.4 used critical values 0.685, 1.079, and 2.098 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$. Source: Authors' elaboration.

Table 5. Forecasting stock market implied volatility with financial reports tone disagreement: Out-of-sample analysis with the ENC-t test.

Contrasting against an AR(1) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \Delta \ln(P_{i,t-1}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	0.452	1.803 **	1.328 *	1.965 **
2	−0.269	1.481 **	1.092 *	1.703 **
1	1.125 *	1.015 *	0.336	1.242 *
0.4	1.358 **	1.729 **	1.604 **	2.128 ***
Contrasting against an AR(2) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^2 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$				
P/R	RVX	VIX	VXD	VXO
4	0.355	1.617 **	1.157 *	1.699 **
2	−0.317	1.319 *	0.999 *	1.485 **
1	1.071 *	0.877	0.214	1.038 *
0.4	1.349 **	1.707 **	1.587 **	2.002 ***

Table 5. Cont.

Contrasting against an AR(3) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^3 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$					
P/R	RVX	VIX	VXD	VXO	
4	0.312	1.546 **	1.050 *	1.618 **	
2	−0.274	1.268 **	0.902	1.422 **	
1	1.022 *	0.865	0.243	1.018 *	
0.4	1.221 *	1.490 **	1.375 *	1.773 **	
Contrasting against an AR(6) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^6 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$					
P/R	RVX	VIX	VXD	VXO	
4	0.545	1.699 **	1.244 *	1.663 **	
2	0.015	1.391 **	1.025 *	1.514 **	
1	1.160 *	0.866	0.391	0.977 *	
0.4	1.417 **	1.406 **	1.480 **	1.741 **	

Notes: Each entry reports the Enc-t statistic with a recursive scheme. This table compares the predictive ability of our out-of-sample core models (See Table 1, row 2) with the predictive ability of the benchmark models (See Table 1, row 3). We report four estimation windows with the following critical values: P/R = 4 used critical values 0.928, 1.332, and 2.120 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 2 used critical values 0.939, 1.322, and 2.082 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 1 used critical values 0.955, 1.331, and 2.052 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. P/R = 0.4 used critical values 1.005, 1.338, and 1.997 for $p < 10\%$, $p < 5\%$, and $p < 1\%$, respectively. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$. Source: Authors' elaboration.

Table 6. Mean Directional Accuracy when Forecasting stock market implied volatility with financial reports tone disagreement.

Contrasting against an AR(3) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^3 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$					
	P/R	RVX	VIX	VXD	VXO
Benchmark	4	−0.023	0.052 **	0.034 *	0.048 **
Over 50%	4	0.517	0.623 ***	0.576 ***	0.605 ***
Benchmark	2	0.007	0.049 **	0.031 *	0.049 **
Over 50%	2	0.542	0.621 ***	0.577 ***	0.598 ***
Benchmark	1	−0.028	0.029	0.041 *	0.035 *
Over 50%	1	0.570 *	0.610 ***	0.579 **	0.583 ***
Benchmark	0.4	0.0322	0.042	0.036	0.063 **
Over 50%	0.4	0.613 ***	0.616 ***	0.510 **	0.613 ***
Contrasting against an AR(6) Benchmark model: $\ln(P_{i,t}) = c + \gamma_i \sum_{n=1}^6 \Delta \ln(P_{i,t-n}) + \varepsilon_{i,t}$					
	P/R	RVX	VIX	VXD	VXO
Benchmark	4	0.017	0.037 *	0.025	0.029 *
Over 50%	4	0.552 *	0.563 ***	0.536 *	0.520 **
Benchmark	2	−0.007	0.040 *	0.025	0.031 *
Over 50%	2	0.550 *	0.562 **	0.546 **	0.554 **
Benchmark	1	−0.018	0.047 *	0.013	0.047 **
Over 50%	1	0.551	0.571 ***	0.544 *	0.571 ***
Benchmark	0.4	0.048 **	0.094 ***	0.012	0.042 **
Over 50%	0.4	0.515 ***	0.513 ***	0.590 ***	0.5100 ***

Notes: Significance levels are established based on Diebold and Mariano [60] and West [62] test statistic (DMW t-stat). Rejecting the null hypothesis means that the core models have a higher success rate than a 50% “pure luck” rate, for the rows titled “Over 50%”. For the rows titled “Benchmark”, rejecting the null hypothesis with positive coefficients means that the core models have a superior success rate than the benchmark models. * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

4.1. In-Sample Analyses

Table 2 contains in-sample results (See Table 1, row 1) of financial reports tone disagreement forecasting for four stock market implied volatility indices: VIX, RVX, VXD, and VXO. Our results show that tone disagreement $\Delta \ln(MSTL_{t-1})$ has the ability to forecast the implied volatility indices. Our results show that tone disagreement can predict the implied volatility indices. Specifically, we find that the R-squared spans from 5.8% to 7.6% with an average of 6.6% across all models. The largest R-squared was yield when forecasting VIX, the most well-known and traded volatility index. In addition, the autoregressive

components were significant for all indices except RVX on the first lag, significant for only VIX and VXO on the second lag, and significant for all indices on the third lag, which was the most significant lag at the 1% significance level across all models.

4.2. In-Sample Analysis with Regime Breaks

Table 3 contains in-sample results based on Table 1, row 1 that are similar to those reported in Table 2 but take into account distinct regime breaks, following Bai and Perron [63]. This method allows the parameters to change dynamically across different time-span scenarios. The R-squared has grown substantially, now ranging from 18.30% to 22.64%, with an average of 20.51%, which is more than three times the average of the R-squared reported in Table 2 (which was 6.6%). The explanation power increase by allowing more dynamic parameter changes across regimes, resulting in a better fit of the models. VIX, VXD, and VXO were significant in four of the six regime breaks, while RVX was significant in two. The findings indicate predictability before, during, and after the 2008 financial crisis.

Interestingly, the coefficients shift from being positive in the regimes before and during the 2008 financial crisis to damaging during the recovery from the 2008 financial crisis (2010M08–2013M12). Furthermore, three of the fourteen significant coefficients had negative coefficients, which all happened during the recovery period. Finally, the 2008 financial crisis does not determine the in-sample results in Table 2 because eleven of the fourteen significant coefficients were positive.

Despite these in-sample indicators, there may still be concerns regarding the actual usefulness of financial reports tone disagreement as a forecasting tool, mainly due to worries regarding data mining and overfitting issues, as highlighted by Clark and McCracken [64]. For this condition, an out-of-sample approach would be more appropriate to determine whether these insights hold true and whether they are helpful to practitioners looking to make real-time predictions, especially those involved in financial stability and portfolio optimization tasks. In the following sections, we will describe a set of results obtained using out-of-sample techniques, allowing for a better understanding of the pervasiveness of the relationship between tone disagreement in financial reports and implied volatility.

4.3. Out-of-Sample Analysis

This subsection presents three out-of-sample exercises: ENCNEW (Table 4), ENC-t (Table 5), and Mean Directional Accuracy Analysis (Table 6). For out-of-sample analyses, we contrast the predictive performance of our out-of-sample core model (See Table 1, row 1) to that of the out-of-sample benchmark model (See Table 1, rows 2 and 3). $H_0 : \beta = 0$ implies that our model simplifies to an $AR_{(p)}$ (Table 1, rows 2 and 3).

The results using just one ad hoc window size, according to Clark and McCracken [64], may still be highly debatable because predictability could only apply to a single subsample and consequently not be resilient to different window sizes. Therefore, to mitigate any concerns about overfitting, we consider four different window sizes ($P/R = 4, 2, 1,$ and 0.4 , which correspond to estimating our model with 20%, 33%, 50%, and 71% of the sample observations, respectively, and evaluating the forecast models with the remaining observations).

Table 4 reports the results from the ENCNEW test, which contrasts the prediction ability of the core models (Table 1, row 1) to the benchmark models (Table 1, rows 2 and 3). Additionally, we added two autoregressive benchmark specifications to extend our analysis: $AR_{(1)}$ and $AR_{(2)}$. The core models outperformed the benchmark models in 72% of the exercises. Remarkably, VIX and VXO could be consistently forecasted, with significant results across all autoregressive and estimation window specifications. On the other hand, models predicting the VXD achieved significant results in 12 out of 16 specifications. Finally, we observe the weakest results predicting RVX, which only yielded a single significant result across all 16 exercises.

The results point to stronger null hypothesis rejections when considering the models that used 71% of the sample to estimate the parameters ($P/R = 0.4$) and the remaining to

evaluate the forecasting models, which revealed significant results in 81% of the exercises. The models with estimation windows specifications of $P/R = 4$ and $P/R = 2$ achieved the same frequency of significant results, with significance found in 75% of the exercises. Finally, the models with estimation windows specifications of $P/R = 1$ achieved the smallest frequency of significance, with significance found in 50% of the exercises. All significant coefficients were consistently positive, consistent with the in-sample predictions that a higher level of financial reports tone disagreement forecasts an increase in implied volatility.

Table 5 reports the results from the ENC-t test, which also contrasts the prediction ability of the core models (Table 1, row 1) to the benchmark models (Table 1, rows 2 and 3) and also considers four different autoregressive benchmark specifications: $AR_{(1)}$, $AR_{(2)}$, $AR_{(3)}$ and $AR_{(6)}$. The core models outperformed the benchmark models in 75% of the exercises. Strikingly, VXO could be consistently forecasted, with significant results across all autoregressive and estimation window specifications. The VIX indices could be forecasted in 14 out of 16 exercises. On the other hand, models predicting the VXD achieved significant results in 11 out of 16 specifications. Again, we obtained the weakest results predicting RVX, which yielded significant results in eight cases across all 16 exercises.

Similarly to findings using the ENCNEW test, we found strong null hypothesis rejections when considering the models that used 71% of the sample to estimate the parameters ($P/R = 0.4$) and the remaining to evaluate the forecasting models, as these models yielded significant results across all exercises. The models with estimation windows specifications of $P/R = 4$ and $P/R = 2$ achieved 75% and 69% of significant results frequency across all specifications. Finally, and similarly to what was also found using the ENCNEW test, the models with estimation windows specifications of $P/R = 1$ achieved the smallest frequency of significance, with significance found in 56% of the exercises. All significant coefficients were positive and consistent with the in-sample and out-of-sample predictions previously discussed.

As a final observation, the core models performed remarkably well predicting the sign change of the two most popular and traded indices, the VIX and VXO. Moreover, they were significantly superior in every exercise, consistent with the ENCNEW and ENC-t test results.

As a last out-of-sample exercise, in Table 6, we report the Mean Directional Accuracy of predictions made using the core models (Table 1, row 1) by calculating the hit rate as a simple average of W_t , as defined in Equation (8), where we contrast the null hypothesis of $H_0 : E(W_t) \leq 0.5$ with the alternative hypothesis $H_1 : E(W_t) > 0.5$, which is a straight comparison against a “pure luck” benchmark. The results in Table 6 are remarkable: As reported on the rows titled “Over 50%”, the core models outperformed the “pure luck” benchmark in 91% of the exercises. Furthermore, considering all exercises, the average hit rate was 14% higher than a 50% “pure luck” benchmark.

The models with estimation windows specifications of $P/R = 0.4$, which use 67% of the sample observations to estimate the parameters and the remaining observations to evaluate the forecast models, yielded significance across all exercises considered. All three remaining estimation windows specifications, $P/R = 4$, 2, and 1, yielded a frequency of 88% of significant results across all exercises. The core models with an $AR_{(6)}$ had a higher frequency of significant results than the core models with an $AR_{(3)}$; the former had 91% of significant results across all exercises, while the latter had a frequency of 88%.

Table 6 also reports on the rows titled “Benchmark,” the difference between the percentage of accurate directional predictions of the core model (Table 1, row 1) and the benchmark model (Table 1, row 2). Considering all exercises, the core models achieved a successful hit rate of 5% higher than the benchmark models, on average. The results considering this approach show that the core models significantly surpass the benchmark models in 56% of the exercises. In addition, all significant coefficients were positive, which also corroborates the superior performance of the core models in predicting the correct sign of change in the volatility indices. The models with estimation windows specifications of $P/R = 4$ and $P/R = 2$ yielded a frequency of significant results equal to 63% across all

exercises, followed by the estimation windows specifications of $P/R = 1$ and $P/R = 0.5$, which yielded both a frequency of significant results equal to 50% across all exercises.

5. Discussion

We study the predictability of implied volatility indices of stocks using financial reports tone disagreement from U.S. firms. Our findings indicate that a higher level of financial reports tone disagreement forecasts an increase in implied volatility. This evidence is consistent with the notion that when the tone varies widely between firms, investors collectively disagree more about the future market outcome, which is plausible to result in increased volatility ultimately.

An interesting discussion around these results is related to Knightian uncertainty [65], in that disagreements among firms could create uncertainty about the market's future or the industry. This theory is in line with previous finds suggesting that a higher level of Knightian uncertainty decreases investment value [66], leading to higher cash holding and a higher probability of derivatives use [67].

We use the term Knightian uncertainty to distinguish risk from uncertainty, where risk refers to randomness that can be measured precisely, such as the probabilities associated with coin tosses. Uncertainty is an inability to perfectly forecast the probability of events occurring, as discussed by Frank Knight in his seminal book *Risk, uncertainty, and profit*.

In this line, an increase in the uncertainty faced by investors could be due to the need for more information or knowledge about the potential outcomes of tone disagreement, making it difficult for market participants to assign probabilities to those outcomes. In this way, the disagreement in the firms' tone could contribute to Knightian uncertainty, leading to increased volatility as market participants try to navigate the uncertain market information setting.

Overall, our findings show that the firm's clear perspective, as expressed by the tone used in financial reports, can provide valuable insights into future market behavior, mainly because the firm's financial reports are an important source of opinion creation for many market participants. Our findings support Atmaz and Basak's [14] hypothesis that a high level of expectation dispersion leads to higher stock volatility.

6. Conclusions

This paper explores the predictability of implied volatility indices of stocks using the level of disagreement in tone between financial reports from U.S. firms. Applying network analysis methods for modeling tone synchronization as a "tone disagreement proxy", we find evidence that a higher tone disagreement predicts a higher variation in implied volatility indices one month ahead. In other words, our results validate the forecasting ability over future volatility in stock markets of tone disagreement regarding the prospect performance of firms.

Our research contributes to the growing literature of disagreement in capital markets that connects how differences in the expectations of financial agents generate particular patterns in financial markets that depart from the classical rational agent models embodied in the traditional asset pricing models.

As the general results of our research show, a greater level of tone disagreement predict an increase in the implied volatility in stock markets, in line with the literature that states that a higher expectations dispersion generated by higher tone disagreement leads to greater stock volatility. Therefore, we suggest that our tone disagreement metric captures the disagreement about the future economic performance, consequently generating higher levels of implied volatility.

Our results relating in a positive way tone disagreement and future implied volatility can help regulators, policymakers, and practitioners supervise the behavior of financial markets and monitor financial risks as well as in the duties derived from the delegated portfolio management role. As the empirical literature suggests, implied volatility indices possess the ability to be a fear gauge of the market agents. Accordingly, being used as a

market trading tool or as a future uncertainty proxy, their influence on the behavior of market agents transforms them into a crucial anticipation of the psychology of the investors.

Our research has some limitations to consider despite its contribution to the empirical literature on tone disagreement. First, we solely analyze firms' reports of U.S. firms. As the U.S. markets are probably the most efficient in the world, the capabilities of financial analysts to process information and incorporate them into stock prices are well-stated. However, in those markets with lessens levels of financial efficiency and development, probably the connection between disagreement and the implied volatility behavior that we found would not be assured. Secondly, following the forecasting literature, we evaluate and test one month ahead prediction of the tone disagreement over the implied volatility indices. Therefore, predictions with longer terms durations would not be extended directly from our results.

Finally, our work brings new research extensions:

1. We assess tone disagreement synchronization among firms' reports by applying the minimum spanning tree (MST) length. However, we do not employ other network analysis metrics and methods to expand our understanding of the network tone, such as centrality, diameter, degree, Planar Maximally Filtered Graph (PMFG).
2. We analyze the behavior of stock market price patterns and tone disagreements. Hence, it is interesting studying the behavior of other financial markets such as commodities, currencies, and bonds to verify if our main results keep when analyzing different financial assets.
3. Our paper uses econometric models combined with text mining estimations to explore our research topic. The usefulness of other empirical methods would be enjoyable to explore.

Author Contributions: Conceptualization, N.S.M., N.H., J.F.L. and T.F.; methodology, N.S.M., N.H. and J.F.L.; software, N.S.M. and N.H.; validation, N.S.M., N.H., J.F.L. and T.F.; formal analysis, N.S.M., N.H., J.F.L. and T.F.; investigation, N.S.M., N.H., J.F.L. and T.F.; resources, N.S.M., N.H., J.F.L. and T.F.; data curation, N.S.M., N.H. and J.F.L.; writing—original draft preparation, N.S.M., N.H., J.F.L. and T.F.; writing—review and editing, N.S.M., N.H., J.F.L. and T.F.; visualization, N.S.M., N.H., J.F.L. and T.F.; supervision, N.S.M., N.H., J.F.L. and T.F.; project administration, N.S.M.; funding acquisition, N.S.M., N.H. and J.F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hong, H.; Stein, J.C. Disagreement and the Stock Market. *J. Econ. Perspect.* **2007**, *21*, 109–128. [[CrossRef](#)]
2. Barberis, N.; Huang, M. Mental Accounting, Loss Aversion, and Individual Stock Returns. *J. Financ.* **2001**, *56*, 1247–1292. [[CrossRef](#)]
3. Barberis, N.; Shleifer, A.; Vishny, R. A Model of Investor Sentiment. *J. Financ.* **1998**, *49*, 307–343.
4. Daniel, K.; Hirshleifer, D.; Subrahmanyam, A. Investor Psychology and Security Market Under- and Overreactions. *J. Financ.* **1998**, *53*, 1839–1885. [[CrossRef](#)]
5. Hirshleifer, D. Investor Psychology and Asset Pricing. *J. Financ.* **2001**, *56*, 1533–1597. [[CrossRef](#)]
6. Barberis, N.; Thaler, R. A Survey of Behavioral Finance. *Handb. Econ. Financ.* **2003**, *1*, 1053–1128. [[CrossRef](#)]
7. Banerjee, S. Learning from Prices and the Dispersion in Beliefs. *Rev. Financ. Stud.* **2011**, *24*, 3025–3068. [[CrossRef](#)]
8. Li, F. The Information Content of Forward-looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *J. Account. Res.* **2010**, *48*, 1049–1102. [[CrossRef](#)]
9. Smith, M.; Taffler, R.J. The Chairman's Statement—A Content Analysis of Discretionary Narrative Disclosures. *Account. Audit. Account. J.* **2000**, *13*, 624–647. [[CrossRef](#)]
10. Davis, A.K.; Tama-Sweet, I. Managers' Use of Language across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. *Contemp. Account. Res.* **2012**, *29*, 804–837.
11. Demers, E.; Vega, C. The Impact of Credibility on the Pricing of Managerial Textual Content. *SSRN* **2014**, 1153450.
12. Huang, X.; Teoh, S.H.; Zhang, Y. Tone Management. *Account. Rev.* **2014**, *89*, 1083–1113. [[CrossRef](#)]

13. Gandhi, P.; Loughran, T.; McDonald, B. Using Annual Report Sentiment as a Proxy for Financial Distress in US Banks. *J. Behav. Financ.* **2019**, *20*, 424–436. [[CrossRef](#)]
14. Atmaz, A.; Basak, S. Belief Dispersion in the Stock Market. *J. Financ.* **2018**, *73*, 1225–1279. [[CrossRef](#)]
15. Scheinkman, J.A.; Xiong, W. Overconfidence and Speculative Bubbles. *J. Political Econ.* **2003**, *111*, 1183–1220. [[CrossRef](#)]
16. Buraschi, A.; Jiltsov, A. Model Uncertainty and Option Markets with Heterogeneous Beliefs. *J. Financ.* **2006**, *61*, 2841–2897. [[CrossRef](#)]
17. Li, T. Heterogeneous Beliefs, Asset Prices, and Volatility in a Pure Exchange Economy. *J. Econ. Dyn. Control* **2007**, *31*, 1697–1727. [[CrossRef](#)]
18. David, A. Heterogeneous Beliefs, Speculation, and the Equity Premium. *J. Financ.* **2008**, *63*, 41–83. [[CrossRef](#)]
19. Dumas, B.; Kurshev, A.; Uppal, R. Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility. *J. Financ.* **2009**, *64*, 579–629. [[CrossRef](#)]
20. Banerjee, S.; Kremer, I. Disagreement and Learning: Dynamic Patterns of Trade. *J. Financ.* **2010**, *65*, 1269–1302. [[CrossRef](#)]
21. Coelho, R.; Gilmore, C.G.; Lucey, B.; Richmond, P.; Hutzler, S. The Evolution of Interdependence in World Equity Markets—Evidence from Minimum Spanning Trees. *Phys. A Stat. Mech. Its Appl.* **2007**, *376*, 455–466. [[CrossRef](#)]
22. Mantegna, R.N. Hierarchical Structure in Financial Markets. *Eur. Phys. J. B* **1999**, *11*, 193–197. [[CrossRef](#)]
23. Fama, E.F.; French, K.R. The Cross-Section of Expected Stock Returns. *J. Financ.* **1992**, *47*, 427–465. [[CrossRef](#)]
24. Pincheira, P.; Hardy, N. Forecasting Aluminum Prices with Commodity Currencies. *Resour. Policy* **2021**, *73*, 102066. [[CrossRef](#)]
25. Clark, T.E.; McCracken, M.W. Tests of Equal Forecast Accuracy for Encompassing for Nested Models. *J. Econom.* **2001**, *15*, 85–110. [[CrossRef](#)]
26. Pincheira-Brown, P.; Bentancor, A.; Hardy, N.; Jarsun, N. Forecasting fuel prices with the Chilean exchange rate: Going beyond the commodity currency hypothesis. *Energy Econ.* **2022**, *106*, 105802. [[CrossRef](#)]
27. Pincheira-Brown, P.; Neumann, F. Can We Beat the Random Walk? The Case of Survey-Based Exchange Rate Forecasts in Chile. *Financ. Res. Lett.* **2020**, *37*, 101380. [[CrossRef](#)]
28. Cheung, Y.-W.; Chinn, M.D.; Pascual, A.G. Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive? *J. Int. Money Financ.* **2005**, *24*, 1150–1175. [[CrossRef](#)]
29. Loughran, T.; McDonald, B. The Use of EDGAR Filings by Investors. *J. Behav. Financ.* **2017**, *18*, 231–248. [[CrossRef](#)]
30. Cohen, L.; Malloy, C.; Nguyen, Q. Lazy Prices. *J. Financ.* **2020**, *75*, 1371–1415. [[CrossRef](#)]
31. Gao, M.; Huang, J. Informing the Market: The Effect of Modern Information Technologies on Information Production. *Rev. Financ. Stud.* **2020**, *33*, 1367–1411. [[CrossRef](#)]
32. Rossi, B.; Inoue, A. Out-of-Sample Forecast Tests Robust to the Choice of Window Size. *J. Bus. Econ. Stat.* **2012**, *30*, 432–453. [[CrossRef](#)]
33. Dzieliński, M.; Hasseltoft, H. News Tone Dispersion and Investor Disagreement. *SSRN* **2017**, 2192532.
34. Cookson, J.A.; Niessner, M. Why Don't We Agree? Evidence from a Social Network of Investors. *J. Financ.* **2020**, *75*, 173–228. [[CrossRef](#)]
35. Feldman, R.; Govindaraj, S.; Livnat, J.; Segal, B. Management's Tone Change, Post Earnings Announcement Drift and Accruals. *Rev. Account. Stud.* **2010**, *15*, 915–953. [[CrossRef](#)]
36. McConnell, D.; Haslem, J.A.; Gibson, V.R. The President's Letter to Stockholders: A New Look. *Financ. Anal. J.* **1986**, *42*, 66–70. [[CrossRef](#)]
37. Swales, G.S., Jr. Another Look at the President's Letter to Stockholders. *Financ. Anal. J.* **1988**, *44*, 71–73. [[CrossRef](#)]
38. Abrahamson, E.; Amir, E. The Information Content of the President's Letter to Shareholders. *J. Bus. Financ. Account.* **1996**, *23*, 1157–1182. [[CrossRef](#)]
39. Kothari, S.P.; Li, X.; Short, J.E. The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *Account. Rev.* **2009**, *84*, 1639–1670. [[CrossRef](#)]
40. Liang, C.; Wei, Y.; Zhang, Y. Is Implied Volatility More Informative for Forecasting Realized Volatility: An International Perspective. *J. Forecast.* **2020**, *39*, 1253–1276. [[CrossRef](#)]
41. Whaley, R.E. The Investor Fear Gauge. *J. Portf. Manag.* **2000**, *26*, 12–17. [[CrossRef](#)]
42. Anatolyev, S.; Gerko, A. A Trading Approach to Testing for Predictability. *J. Bus. Econ. Stat.* **2005**, *23*, 455–461. [[CrossRef](#)]
43. Kang, S.H.; Maitra, D.; Dash, S.R.; Brooks, R. Dynamic Spillovers and Connectedness between Stock, Commodities, Bonds, and VIX Markets. *Pac. Basin Financ. J.* **2019**, *58*, 101221. [[CrossRef](#)]
44. Shu, H.C.; Chang, J.H. Spillovers of Volatility Index: Evidence from U.S., European, and Asian Stock Markets. *Appl. Econ.* **2019**, *51*, 2070–2083. [[CrossRef](#)]
45. Loughran, T.; McDonald, B. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *J. Financ.* **2011**, *66*, 35–65. [[CrossRef](#)]
46. Tumminello, M.; Lillo, F.; Mantegna, R.N. Correlation, hierarchies, and networks in financial markets. *J. Econ. Behav. Organ.* **2010**, *75*, 40–58. [[CrossRef](#)]
47. Magner, N.S.; Lavin, J.F.; Valle, M.A.; Hardy, N. The Volatility Forecasting Power of Financial Network Analysis. *Complexity* **2020**, *33*, 916–922. [[CrossRef](#)]
48. Lavin, J.F.; Valle, M.A.; Magner, N.S. A Network-Based Approach to Study Returns Synchronization of Stocks: The Case of Global Equity Markets. *Complexity* **2021**, *2021*, 7676457. [[CrossRef](#)]

49. Magner, N.; Lavin, J.F.; Valle, M.; Hardy, N. The Predictive Power of Stock Market's Expectations Volatility: A Financial Synchronization Phenomenon. *PLoS ONE* **2021**, *16*, e0250846. [[CrossRef](#)]
50. Yan, X.-G.; Xie, C.; Wang, G.-J. Stock Market Network's Topological Stability: Evidence from Planar Maximally Filtered Graph and Minimal Spanning Tree. *Int. J. Mod. Phys. B* **2015**, *29*, 1550161. [[CrossRef](#)]
51. Wen, F.; Yuan, Y.; Zhou, W. Cross-shareholding Networks and Stock Price Synchronicity: Evidence from China. *Int. J. Financ. Econ.* **2021**, *26*, 914–948. [[CrossRef](#)]
52. Onnela, J.P.; Chakraborti, A.; Kaski, K.; Kertész, J.; Kanto, A. Dynamics of Market Correlations: Taxonomy and Portfolio Analysis. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **2003**, *68*, 056110. [[CrossRef](#)]
53. Newey, W.K.; West, K.D. Hypothesis Testing with Efficient Method of Moments Estimation. *Int. Econ. Rev.* **1987**, *28*, 777–787. [[CrossRef](#)]
54. Newey, W.K.; West, K.D. Automatic Lag Selection in Covariance Matrix Estimation. *Rev. Econ. Stud.* **1994**, *61*, 631–653. [[CrossRef](#)]
55. Timmermann, A. Elusive Return Predictability. *Int. J. Forecast.* **2008**, *24*, 1–18. [[CrossRef](#)]
56. Pincheira, P.; Hardy, N. Forecasting Base Metal Prices with the Chilean Exchange Rate. *Resour. Policy* **2019**, *62*, 256–281. [[CrossRef](#)]
57. Bai, Y. Country Factors in Stock Returns: Reconsidering the Basic Method. *Appl. Financ. Econ.* **2014**, *24*, 871–888. [[CrossRef](#)]
58. Clark, T.E.; West, K.D. Approximately normal tests for equal predictive accuracy in nested models. *J. Econom.* **2007**, *138*, 291–311. [[CrossRef](#)]
59. Clark, T.E.; McCracken, M.W. Nested forecast model comparisons: A new approach to testing equal accuracy. *J. Econom.* **2015**, *186*, 160–177. [[CrossRef](#)]
60. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *20*, 134–144. [[CrossRef](#)]
61. Pesaran, M.H.; Timmermann, A. Predictability of Stock Returns: Robustness and Economic Significance. *J. Financ.* **1995**, *50*, 1201–1228. [[CrossRef](#)]
62. West, K.D. Asymptotic Inference about Predictive Ability. *Econometrica* **1996**, *64*, 1067–1084. [[CrossRef](#)]
63. Bai, J.; Perron, P. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* **1998**, *66*, 47–78. [[CrossRef](#)]
64. Clark, T.; McCracken, M. Advances in forecast evaluation. *Handb. Econ. Forecast.* **2013**, *2*, 1107–1201.
65. Knight, F.H. *Risk, Uncertainty and Profit*; Houghton Mifflin: Boston, MA, USA; New York, NY, USA, 1921; Volume 31.
66. Nishimura, K.G.; Ozaki, H. Irreversible Investment and Knightian Uncertainty. *J. Econ. Theory* **2007**, *136*, 668–694. [[CrossRef](#)]
67. Friberg, R.; Seiler, T. Risk and Ambiguity in 10-Ks: An Examination of Cash Holding and Derivatives Use. *J. Corp. Financ.* **2017**, *45*, 608–631. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.