Investor Sentiment and Volume-Volatility Relationship

Yuan Xue†

October 5, 2017

Abstract
This paper shows the effect of investor sentiment on information processing in the financial market. We investigate how disagreement among investors affects the relationship between trading intensity and price volatility around macroeconomic announcements during high and low sentiment periods. By incorporating into the Kandel and Pearson (1995) model a one factor structure with heterogeneous beliefs in the idiosyncratic components, we explicitly derive the volume-volatility elasticity for individual stocks around systematic information release. Our empirical results are based on intraday transaction data of S&P 500 ETF and Dow Jones 30 components, as well as high frequency econometric tools for the multi-dimensional setting. Consistent with the model predictions, our estimates of elasticity decrease significantly with the ratio of idiosyncratic variance. Disagreement measures only cast significantly negative effect in high sentiment periods for both the market portfolio and individual stocks, which is in line with changes in investors’ confidence level when sentiment regime shifts.

Keywords: Differences-of-opinion; investor sentiment; high-frequency data; trading volume; stochastic volatility; economic uncertainty.

JEL classification: C52, G12, G41.

*I am grateful my advisors Jia Li and Tim Bollerslev for their invaluable help in producing this paper. I would also like to thank George Tauchen, Federico Bugni, Andrew Patton, and Brian Weller as well as all the participants of Dukes financial econometrics work group and participants at other seminars and conferences for their helpful comments and suggestions. All errors are mine.

†Department of Economics, Duke University, Durham, NC 27708; e-mail: yuan.xue@duke.edu.
1 Introduction

How information is processed in the financial market is of high interest. In literature, comovement of volume and volatility upon information arrival has been widely documented in the literature. Empirically, researchers find large trading volume and return volatility around news release. Some theoretical models to characterize this processes are the “difference-of-opinion” models, in which investors agree to disagree on common public signals (See Harris and Raviv (1993), Kandel and Pearson (1995), Scheinkman and Xiong (2003) and Banerjee and Kremer (2010)). In Bollerslev et al. (2017), the authors show the volume and volatility elasticity derived from Kandel and Pearson (1995) significantly decreases in the level of disagreement. We propose econometric tools to estimate the volatility-volume elasticity for the S&P 500 ETF, and the empirical results are consistent with the model predictions.

In these “difference-of-opinion” models, investors interpret signals and change positions based on Bayesian learning, and this process can be subject to investor sentiment. Psychology literature proposes that emotions can distort how people process information and make decisions. Tiedens and Linton (2001) cites a large literature that associates hope, surprise, fear, worry, and sadness to the sense of uncertainty. This distortion of behavior in financial market is generally believed to contribute to price anomalies in financial markets, such as overreaction and underreaction to news.(See Bondt and Thaler (1985), De Bondt and Thaler (1987), Bernard and Thomas (1990) Jegadeesh and Titman (1993)). One of the most influential and thus most thoroughly studied emotion in the financial market is investor sentiment, as shown in Lee et al. (1991), Barberis et al. (1998) and Baker and Wurgler (2006). Using proxies for sentiment, researchers find disparity in investor behavior and asset prices during high and low sentiment periods.(See Bergman and Roychowdhury (2008), Antoniou et al. (2013), Yu and Yuan (2011), Stambaugh et al. (2012)). Considering that volume-volatility relationship is the result of investors’ interaction around information release, we hypothesize that it will act differently when sentiment level changes.

To this end, we separate our sample into high and low sentiment periods based on the sample median of the sentiment index from Baker and Wurgler (2006). Following BLX, we estimate the effect of disagreement measures on volume-volatility elasticity in both the high and low sentiment periods using high frequency data. The estimate is based on jumps in nonparametrically estimated volume intensity and spot volatility upon information arrival, which follows the idea of regression

\footnote{See Chaboud et al. (2008), Boudt and Petitjean (2014), Jiang et al. (2011), and Crego (2017)}
discontinuity design as in Hahn et al. (2001). We also implement difference-in-difference to control for intraday seasonality. Disagreement measures include dispersion on real economy, economic uncertainty level, as well as a text based measure of negative tone. The subsample analysis reveals that the negative effect of the disagreement measures are only significant during high sentiment periods for the S&P 500 ETF. In sharp contrast of this, none of the disagreement measures have significant explanatory power in low sentiment periods.

This result is consistent with theoretical predictions. In the difference-of-opinions models, the effect of investor disagreement is magnified by investors’ confidence in interpreting the information from public news. In our sample, low sentiment periods mainly correspond to recessions, so the diminished explanatory power of disagreement measures demonstrates decrease in investors' confidence during recessions. There can be several underlying reasons for this change. Psychologically, people become regretful and lose confidence when they suffer from losses in a bearish market. Indeed, in finance literature, confidence and sentiment are closely related, and in some papers researchers directly use index of confidence to proxy for sentiment (See Bergman and Roychowdhury (2008) and Bergman and Roychowdhury (2008)). Besides, according to Yu and Yuan (2011), more sentiment investors enter the market when sentiment is high, and they are more prone to be less rational and naive. In other words, these investors are more likely to become victim of overconfidence, and their overconfidence can be further aggravated by self-attribution. The combined effect of these channels contribute to the comovement of sentiment and confidence level. As a result, when sentiment is low, even though dispersion in investors’ interpretation still exists, they no longer cast effect in the volume-volatility elasticity because investors themselves do not take these interpretations seriously.

We go one step further to investigate whether the same pattern holds for individual stocks. Individual stocks are both closely related to and different from the market ETF. Apart from the systematic component, they also consist of idiosyncratic components. Though in classic asset pricing models idiosyncratic risks are seldom priced, they can still introduce additional risks when portfolio is perfectly diversified, which are priced and have predictive power as shown in Merton (1987) and Hahn et al. (2001). Empirically though, there is no consensus on how idiosyncratic risks affect prices. See Ang et al. (2006), Fu (2009), and Huang et al. (2009) for some inconclusive results. In line with this strand of literature and the difference-of-opinion models, we assume that investors associate the idiosyncratic component of individual stocks with heterogeneous prices that

\footnote{Our aim is to focus only on the "abnormal" change in the volume intensity and spot volatility around information release, which is absent of trading behaviors not triggered by news, such as noisy trading and low frequency dynamics. Therefore, though the methodology is similar to regression discontinuity design, we do not intend to study the treatment effect as most RDD literature does.}
can be different from zero.

We incorporate a one factor model with disagreement on the idiosyncratic term to extend the Kandel and Pearson (1995) model to a multi-asset setting, which admits the explic solution of the volume-volatility elasticity for individual stocks. We show theoretically that the elasticity decreases monotonically in the overall relative disagreement level, which can further be decomposed into investors’ relative disagreement on the interpretation of public signals and on the idiosyncratic component. The model demonstrates that for both the market index and individual stocks, disagreement measures only matter when they have high confidence in interpreting the public signals. Besides, the higher the ratio of idiosyncratic volatility is, the more disagreement can be introduced by the firm specific term.

To investigate the elasticities cross-sectionally, we adapt the previously mentioned econometric methods to a panel version. Using a fixed effect panel regression with DID, we estimate the elasticity of individual stocks around macroeconomic announcements using the Dow Jones 30 components. We then propose a multi-dimensional bootstrap to facilitate feasible statistical inferences. Since the relative disagreement measure is a dispersion on common signals relative to private information asymmetry, we modify our relative disagreement measure as the difference between standardized disagreement measures and analysts’ forecast dispersion. The idiosyncratic ratio is the ratio between the idiosyncratic volatility to the systematic volatility, estimated using high frequency data. Consistent with the increases in perceived complexity in pricing individual stocks, the elasticities are much lower than those of S&P 500 ETF. In accordance with the model predictions, idiosyncratic ratio decreases elasticity with both statistical and economic significance. This is quite intuitive. Indeed, due to heterogeneity in idiosyncratic risks, it is harder for investors to reach a consensus on the high-idiosyncratic-volatility stocks. The relative disagreement measures also negatively affect the elasticity. Importantly, the result is still mainly driven by the high sentiment periods, the same as for the S&P 500 ETF.

Finally, we implement the same analysis on treasury bond futures as a benchmark. We find that the negative effect of disagreement measures is significant in both high and low sentiment period, and the differences of the coefficients are much weaker than for the securities. This result is as expected, since all the sentiment proxies used in the PCA in Baker and Wurgler (2006) measures the speculative intent in the stock market, so it may not be able to capture the sentiment level in the treasury bond market. Besides, the disparity in the effect of disagreement measures depends on the diminished confidence level in interpreting the assets’ payoff from public news while traders in the bond market may not perceive the same confidence loss during low sentiment periods. Indeed,
the "flight to quality" or "flight to liquidity" phenomena are widely documented in literature, which indicates relatively high confidence for participants in the treasury market during recessions (See Bernanke et al. [1994], Scholes [2000], Longstaff [2002], Lang and Nakamura [1995] and Caballero and Kurlat [2008]).

For robustness check, we do the analysis for redefined high and low sentiment periods using different cutoff points. We also redo the analysis using other disagreement measures. This essentially makes no conclusion to our main conclusion.

This paper contributes to the literature of sentiment. Daniel et al. [1998] and Barberis et al. [1998] propose models with investor sentiment, confidence and self-attritoin to address overreaction and underreaction. Our result is consistent with change in confidence level in different sentiment periods, and thus adds evidences to these models. Researchers have also documented different investor behavior during high and low sentiment periods. For example, Bergman and Roychowdhury [2008] show that managers intentionally increase disclosures of earnings forecasts during low sentiment period to maintain an optimistic profile of the corporate. There are even more evidences of difference in prices when sentiment regime shifts. Antoniou et al. [2013] find that momentum strategy is only profitable when sentiment is optimistic, while Yu and Yuan [2011] show that the positive mean-variance tradeoff and the negative correlation between returns and contemporaneous volatility innovations only exists in low sentiment periods. The disparity of disagreement measures’ effect presented in this paper is consistent with this literature. However, instead of studying prices, we zoom into the short window around information releases and learn investors’ trading behavior through study on volume-volatility relationship. Therefore, we show that sentiment can cast effect on the financial market by affecting investors’ decision making upon information arrival.

This paper also makes a contribution to the literature of volume and volatility. For individual stocks, though there are papers on price change, trading intensity and volatility of individual stocks (See Lamoureux and Lastrapes [1990], Darrat et al. [2007] and Giot et al. [2010] for example), our paper is the first that combines theoretical model with empirical practices based on high frequency data. We extend the difference-of-opinion model and provide theoretical foundation for the volume-volatility relationship around information release. Motivated by the model predictions, we apply econometric methods adapted to a multi-dimensional setting and present the empirical results, which adds new evidences to information processing upon systematic news release for individual stocks.

A third contribution of this paper is to the literature of idiosyncratic risks. In line with many papers in the literature, we assume in our model that the idiosyncratic term is priced. However,
empirically, instead of focusing on the price and idiosyncratic volatility relationship, we show that
the stock with larger idiosyncratic term has smaller volume-volatility elasticity. This result is
consistent with the idea that larger idiosyncratic term introduces more dispersion in beliefs about
the stock’s payoff. In this way, we show that the idiosyncratic term plays a role in the financial
market from another aspective.

This paper precedes as follows. Section 2 presents the model that motivates our empirical
implementations. Section 3 describes the data we use for this paper, followed by the econometric
methodologies in Section 4. The empirical results are shown in Section 5. Section 6 assesses the
robustness of the findings. Section 7 concludes.

2 Theoretical motivation

Our model incorporate a one factor structure with heterogeneous beliefs into Kandel and Pearson
(1995) so that we can explicitly derive an individual stock’s volume-volatility elasticity around
macroeconomic announcements. Since all these macroeconomic announcements are systematic
news, it is natural to decompose the payoff of individual stock to a systematic component and an
idiosyncratic one. Following the setting of Kandel and Pearson (1995), we impose heterogeneous
beliefs in both components. We present the main assumptions and model predictions in this section,
and detailed derivation can be found in Appendix A.2.

2.1 A one factor structure with heterogeneous beliefs

Suppose in the market there are two categories of assets: a risk free asset with zero return rate and
$n$ risky securities with an uncertain payoff $V$ where $V = [V_1 \ V_2 \ \ldots \ V_n]$. There are two groups
of investors among whom a proportion of $\alpha$ are of type 1.

Assumption 1. Investor $j$ believes the payoff of asset $i$ is

$$V_{ij} = \beta_i V_{mj} + \varepsilon_{ij}$$

(2.1)

where $V_{mj} \sim \mathcal{N}(x_{mj}, z_{mj}^2)$, and $\varepsilon_{ij} \sim \mathcal{N}(\gamma_i \eta_j, \gamma_i^2 z_{mj}^2)$.

In the one factor structure, the market component follows a normal distribution with mean $x_{mj}$
and $z_{mj}^2$, which is standard in the one factor model. Investors disagree on the mean and variance
of the market payoff.

Notice that the idiosyncratic term of stock $i$ is $\gamma_i z_{mj}^2$, where the idiosyncratic coefficient $\gamma_i$
measures the magnitude of idiosyncratic volatility relative to the market volatility. Therefore, $\frac{\gamma_i}{\beta_i}$
is the ratio of idiosyncratic volatility to systematic volatility for each stock. The larger the ratio is, the more distinct this stock is from the market index.

For investor $j$, the idiosyncratic term has mean $\gamma_i \eta_j$ instead of 0. This heterogeneous belief may be due to private information, expertise, or simply biased belief in the particular stock $i$. It is magnified by $\gamma_i$, with the intuition that stocks with larger idiosyncratic component tend to generate high heterogeneity in beliefs. This assumption is in line with Stambaugh et al. (2015), which documents a positive relationship between idiosyncratic volatility and mispricing. Similar to the setting in their paper, $\eta_j$ can be positive, negative or zero. However, Our result does not depend on the sign of $\eta_j$ but only requires that $\eta_j$s be different: one group is relatively optimistic and another relatively pessimistic. In our assumption $\eta_j$ can aggregate to 0, which means we do not rely on mispricing to generate our result.

In conclusion, the assumption states that these investors agree on $\beta_i$ and $\gamma_i$, which means they agree on the decomposition of the variance of each individual stock. However, they disagree on the mean and volatility of the systematic component. They also disagree on the mean of unit idiosyncratic term, $\eta_i$.

Admittedly, investors can also disagree on $\beta_i$ and $\gamma_i$, which means investors disagree on the relative magnitude of the systematic component and idiosyncratic component. Consider that we are to study how different interpretations of systematic news affect the volume-volatility elasticity, we do not consider it necessary to introduce different beliefs on the two parameters, which will make the model intractable.

### 2.2 The volume-volatility elasticity

Following the same setting as Kandel and Pearson (1995), after we solve for the equilibrium prices in both the two stages, the equilibrium price change $\Delta P^*$ has the following expression:

$$\Delta P^* = \beta F_1 (L) + \gamma \mu_{\eta}$$

Actually, $F_1 (L)$ has exactly the same expression as the price change in Kandel and Pearson (1995), which can be interpreted as the price change of the market price after the public signal is released. $\mu_{\eta} \propto \eta_1 - \eta_2$ is a constant, the magnitude of which is small compared to that of $\sigma_{pm}$, under certain constraints of the parameters. Therefore, $\Delta P^*$ almost follows a perfect one factor structure around news announcement.

We assume that $F_1 \sim (0, \sigma_{pm}^2)$ where $\sigma_{pm}$ is the volatility of the market price change. On average the price jump is nearly zero for high frequency data, so the mean is set to 0. The volatility of each stock $i$ is then $\sigma_{pi} = \beta_i \sigma_{pm}$, and we can correspondingly derive the joint distribution of the
price changes.

$$\Delta \mathbf{P}^* \sim N(\gamma \mu_{ij}, \beta \gamma \sigma_{pm}^2)$$

With these assumptions, the equilibrium volume $$\text{Vol}_i^*$$ is a function of the price change $$\Delta \mathbf{P}^*$$. Both of the volume and price change are random, and empirically we observe the volume and price change in an "on average" sense. That is, we apply the moment condition and further solve for the relationship between volume intensity $$v_i$$, the average volume, and return volatility $$\sigma_{pi}$$, the average price change for stock $$i$$. The elasticity of $$v_i$$ on $$\sigma_{pi}$$ is

$$\mathcal{E}_i \equiv \frac{\partial v_i(\sigma_{pi})/v_i(\sigma_{pi})}{\partial \sigma_{pi}/\sigma_{pi}} = \frac{1}{1 + \psi(\theta_i)} \left(1 + \frac{\psi(\theta_i)}{2}\right)$$

where $$\psi(x) = x \left[ \frac{\Phi(x)}{\phi(x)} - \frac{1}{2} \right] > 0$$ increases monotonously when $$x > 0$$. We define $$\theta_i$$ as the "total disagreement" measure. For simplicity, we further assume $$\sigma_{m1}^2 = \sigma_{m2}^2 = h^{-1}$$, then $$h$$ can be viewed as the confidence in interpreting the signals. Correspondingly, $$s_j = z_{mj}^2$$ donotes the precision of the prior belief in the market payoff, then the total disagreement $$\theta_i$$ can be further decomposed as follows:

$$\theta_i = \theta_{i \text{sys}} + \theta_{i \text{idio}}$$

$$\theta_{i \text{sys}} = \frac{h}{|s_1-s_2|} \frac{\Delta \mu_m}{\sigma_{pm}}$$

$$\theta_{i \text{idio}} = \left(\frac{h}{|s_1-s_2|} \xi(h_{s_1}, s_2)\right) g\left(\frac{\gamma_i}{\beta_i}\right) \frac{\Delta \eta}{\sigma_{pm}}$$

In the idiosyncratic disagreement component, $$\xi(h, s_1, s_2)$$ is a positive function as long as $$s_1 - s_2 \neq 0$$. $$g\left(\frac{\gamma_i}{\beta_i}\right)_{i}$$ is an increasing function of $$\frac{\gamma_i}{\beta_i}$$. $$\Delta \mu_m$$ and $$\Delta \eta$$ characterize investors' heterogeneous beliefs in the payoff of the interpreted market term and the (per unit) idiosyncratic term:

$$\Delta \mu_m = \mu_{m2} - \mu_{m1}, \quad \Delta \eta = \eta_1 - \eta_2.$$
2.3 Testable hypotheses of the model

When $\Delta \eta = 0$, the elasticity is exactly the market elasticity. $\sigma_{p_m}$ normalizes $\Delta \mu_m$, and we define $\Delta (\mu_m \sigma_{p_m} | s_1 - s_2 |)^{-1}$ as the relative disagreement. It measures the disagreement on the same signal relative to private information before the signal is released. When $\Delta \eta > 0$. $g \left( \frac{s_1}{\bar{s}} \right)$ monotonically increases in $\frac{s_1}{\bar{s}}$. Recall that $E_i$ is a decreasing function of $\theta_i$, we have the following prediction:

Claim 1. When $h$ is large, that is when investors have high confidence in interpreting the public signals.

1. The relative disagreement measure negatively affects the elasticity for both the market ETF and individual stocks.
2. $\frac{s_1}{\bar{s}}$ decreases elasticity.

Claim 2. When $h$ is small, that is when investors have low confidence in interpreting the public signals.

1. The relative disagreement measure has no effect on the elasticity for both the market ETF and individual stocks.
2. $\frac{s_1}{\bar{s}}$ still decreases elasticity.

We will test them in Section 4, but before turning to that we present the data we use.

3 Data description and summary statistics

In this section, we describe the data we use, which include the one minute transaction data, the macroeconomic announcements, the sentiment index, as well as the explanatory variables used in the econometric models.

3.1 High-frequency market prices and trading volume

Our individual stock sample is comprised of Dow Jones 30 Components, as updated in March 18th, 2015. Three stocks (CVX, TRV and V) are excluded because of large period of missing data during the entire sample, from April 10th, 2001 to December 30th, 2014. The S&P 500 index ETF (ticker: SPY) is used as market index, both in regressions for market index as well as when beta and idiosyncratic variances are calculated. These one minute prices and volume data are obtained from TAQ database, cleaned following Brownlees and Gallo (2006) and Barndorff-Nielsen et al.
We also drop all half trading days following previous literature practices, and this results in 1,340,430 one-minute return and volume observations spanning 3437 trading days. The transaction data of ten year treasury bonds are from TickData, and span from July 1, 2003 to September 30, 2014.

3.2 Macroeconomic news announcements

Following BLX, we use the macroeconomic news announcements from the Economic Calendar Economic Release section in Bloomberg. Since individual stocks do not react as strongly as the market index, we only keep the most important indexes in BLX, plus housing index. The important indexes include FOMC announcements, ISM manufacture index, ISM nonmanufacture index, as well as Consumer confidence index. Housing index is composed of all news announcements related to housing, including Existing Home Index, Housing Price Index, NAHB’s HMI, New Home Sales Index and Pending Home Sale Index. Therefore, the housing index happen to include most important indicators in the housing market. In this way, our results are less prone to be contaminated by random noises. We report the sample size and the typical announcement time in Table I. Since there are coincidences of events at the same time, our sample contains 1116 events in total, from April 10th, 2001 to December 31st, 2014.

Table 1: Categories of macroeconomic news announcements

<table>
<thead>
<tr>
<th>Number</th>
<th>Time</th>
<th>Release Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOMC Rate Decision</td>
<td>111</td>
<td>14:15†</td>
</tr>
<tr>
<td>ISM Manufacturing</td>
<td>162</td>
<td>10:00</td>
</tr>
<tr>
<td>ISM Non-Manf. Composite</td>
<td>161</td>
<td>10:00</td>
</tr>
<tr>
<td>Consumer Confidence Index</td>
<td>163</td>
<td>10:00</td>
</tr>
<tr>
<td>Housing Index§</td>
<td>607</td>
<td>10:00 and 13:00</td>
</tr>
</tbody>
</table>

† Exact time varies from 14:00 to 14:15.
§ Includes Existing Home Index, Housing Price Index, NAHB’s HMI, New Home Sales Index and Pending Home Sale Index.

The table presents the total number of announcements for each of the different indicators over the April 10, 2001 to December 31, 2014 sample, along with the release times, and the source of the announcements.

3.3 The Sentiment Measure

The sentiment measure is from Baker and Wurgler (2006). It is based on the first principal component of five sentiment proxies: the closed-end fund discount, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium. This measure of

³There are six proxies in the original paper, while NYSE share turnover is dropped since “turnover does not mean what it once did, given the explosion of institutional high-frequency trading and the migration of trading to a variety of venues”, as described in their data file.
sentiment fits better into our story than the classic consumer confidence type of indexes because it captures the comovement of several sentiment proxies in the financial market with the principal component analysis method. In our paper, we zoom into the one hour window around each macroeconomic announcement and estimate the abnormal volume and volatility, which is the net increase and is more or less clean of noise trading already. Therefore, we care more about the sentiment of the professional financial market participants rather than that of household, though the two may be closely correlated. Following Stambaugh et al. (2012) and Yu and Yuan (2011), we separate our sample into two different subsamples: high sentiment periods and low sentiment periods depending on whether the sentiment measure is above (or equal to) or below the sample median.

Figure 1: The Sentiment Measure

Notes: The figure plots the monthly sentiment measure, the daily annualized realized volatility of S&P 500 ETF, as well as the daily volume. The shaded area is the period of high sentiment.

Figure 1 shows the plot of the sentiment measure along with the daily annualized realized volatility and daily volume. From April 10th, 2001 to December 31st, 2014, there are two concentrated periods of high sentiment: the Internet bubble and the dramatic economy expansion before the subprime mortgage crisis. In the latter half of the sample there are some high sentiment months scattering around the European debt crisis and near the end of our sample. On the other hand, the low sentiment period mainly correspond to economic recessions after market crashes.

The realized volatility and volume are highly correlated, especially during the latter half of the
sample when high frequency trading pervades. There are several peaks of realized volatility: around 2002, in the latter half of 2008 and around 2011. The sentiment index and volatility do not comove throughout the whole sample, though. During the Internet bubble burst, sentiment is high and volatility is also high. However, prior to the subprime mortgage crisis, the expansion in economy is characterized by low volatility and high sentiment. Consider that volatility is highly correlated to risk measures, the sentiment measure is distinctive from those traditional risk measures.

3.4 Explanatory Variables: Disagreement Measures

We use the dispersion in one year ahead unemployment rate and the real GDP growth (in the robustness test) to proxy for difference in opinion in the systematic component of the stock market, since they are of crucial importance in real economy activities. For the treasury bond futures, one year ahead and ten year ahead forecast dispersion in CPI (in the robustness test) are added to account for CPI’s significant impact in the bond market. To avoid peeping into the future, the previous quarter’s value is used as explanatory variable for the current quarter.

Consider the documented high correlation of uncertainty measures with disagreement, we also introduce some uncertainty measures. One measure is NVIX, the news implied volatility measure from [Manela and Moreira](2017). This measure, though aiming to predict VIX, differs from VIX in that it applies textual analysis and machine learning techniques on the title and abstract of all front-page articles of the Wall Street Journal and extracts features to predict VIX. In this way, it avoids direct measure of variation in VIX, which may be correlated to volatility jump, our regressor. This measure is positively correlated with the Unemployment (RGDP growth) rate dispersion, but it has higher variation and the correlation coefficient is 0.398 (0.396). It is a news based measure and thus captures some different aspects from the RGDP dispersion.

Another uncertainty measure is Monthly Index from [Baker et al.](2016). Compared to the weekly policy measure in BLX which is the a weekly moving average of the Daily Policy, there are several reasons for this change. Firstly, this measure has three components: newspaper coverage of policy-related economic uncertainty, the number of federal tax code provisions set to expire in future years, as well as the dispersion in Economic forecast dispersion. The third component includes the dispersions in three indexes: consumer price index (CPI), purchase of goods and services by state and local governments, and purchases of goods and services by the federal government. As stated on the data’s webpage[^4] macroeconomic measures are directly influenced by monetary policy and fiscal policy actions, and thus can be a proxy for uncertainty directly from these policy changes.

[^4]: A full description of the data can be found on the [data website](#)
This measure also accounts for the uncertainty about the tax code, which is, as stated by the authors, a source of uncertainty for businesses and households. Therefore, this new hybrid measure is more comprehensive, which should be more suitable when we include individual stocks. A second reason for using this measure is that it is of lower frequency, and thus less noisy than Weekly Policy. Consider the increased noise introduced by estimation of the volume and volatility of individual stocks, we may prefer a uncertainty measure with smaller noises.

We also add into elasticity the FOMC sentiment measure from BLX. It is is a measure of how negative the tone of each FOMC statements are, constructed following Loughran and McDonald (2011). It is similar to the textual measures mentioned above, but may have higher explanatory power for FOMC announcements since it is directly related to the contents of the announcements.

Figure 2 plots the explanatory series. We first look at the low disagreement periods: before the financial crisis, high sentiment is accompanied by low disagreement, while after European debt crisis, both sentiment and dispersion are low. There are also two peaks of dispersions: the burst of Internet bubble is characterized by high dispersion and high sentiment, while the subprime mortgage crises sees high dispersion and low sentiment. Therefore, our sample includes various high-low combinations of dispersion and sentiment measures.

Generally, the uncertainty measures are very similar in trend to the dispersion measures. The most obvious difference is that NVIX tends to stay higher than the dispersion measures after the subprime mortgage crisis. That means, in the second economy expansion stage, the sentiment and dispersion is generally low, but uncertainty remains relatively high.

The summary statistics of these explanatory variables are presented in Table 2. Each time series has been normalized to have a standard deviation of 1. Consistent with Kim et al. (2014), we find that the summary statistics of the disagreement measures are similar in high and low sentiment periods, but they seem to be slightly higher when sentiment is low. Similarly, the uncertainty measures are also higher with larger standard deviation when sentiment is low. In Bloom (2014), the author finds that uncertainty rises sharply during recessions, so this table is consistent with his research, considering that a majority of low sentiment periods is during recessions. The negative measure stays low during the first half of our sample and gradually climbs up after 2007 before dropping down at the beginning of 2014. As the sentiment is low in the latter half of our sample, low sentiment period cooccurs with a higher negative measure on average.
Notes: The figure plots the one quarter ahead forecast dispersion in unemployment rate and CPI, t, NVIX, MonthlyIndex and the Negative measure throughout the whole sample. The top panel plots the daily value of the whole sample while the bottom one only includes days around FOMC announcements.

3.5 Explanatory Variables: Cross-sectional Variables

Asymmetric information plays a larger role in individual stock’s valuation. For the market index and the treasury bonds, it is hard for investors to possess private information about the whole financial market. Even if investors have access to private information about individual stocks, these different sources of information will aggregate out so they will not affect the pricing of the whole market. However, for individual stocks, there is extended literature in support of the existence of private information. Mayew et al. (2013) show that analysts who ask questions in conference calls possess private information advantages. Analysts’ earnings forecasts for firms composed of more intangibles are documented to contain higher proportions of private information in Barron et al. (2002). Actually, it has been recognized in literature that analysts make forecasts both on publicly available and privately acquired information (e.g., Arnold and Moizer (1984), Barker (1998)). Therefore, it is reasonable to assume the existence of variations in information asymmetry among investors, both cross-sectionally and over time.

One common proxy for private information is analysts’ dispersion. In this paper, analysts’ dispersion is defined as the monthly standard deviation of one year ahead forecasts for EPS, nor-
The table reports the summary statistics of the explanatory variables. The mean, standard deviation, median, minimum, 25% quantile, 75% quantile as well as the maximum of the explanatory variables are calculated for the whole sample, and the high(low) sentiment subsample. Each time series has been normalized so that the standard deviation is 1 for the whole sample.

malized by the average price of the same month. The standard deviation is from I/B/E/S while the prices are from Compustat. To avoid peeping into future, we use the previous month’s analysts’ dispersion for the current month.

The idiosyncratic ratio defined in section can be approximated using high frequency stock returns. We follow [Patton and Verardo (2012)] and use 15 minutes’ return to obtain monthly beta, $\beta_{it}$ for each individual stock. The monthly idiosyncratic variances are calculated as follows:

$$\text{IdioRV}_{it} = \text{RV}_{it} - \beta_{it}^2 \text{RV}_{mt}$$

where $\text{RV}_{it}$ is the monthly realized variance of stock $i$, calculated using 15 minutes’ return again. $\text{RV}_{mt}$ is the concurrent market realized variance. To be consistent with the topic in this paper, we only include intraday returns. We define Idio-Sys Ratio as the ratio of idiosyncratic variance to systematic variance:

$$\text{Idio-Sys Ratio} = \sqrt{\frac{\text{IdioRV}_{it}}{\beta_{it}^2 \text{RV}_{mt}}}$$

(3.1)

Figure 3 shows the 95% empirical confidence interval and the mean value of analysts’ dispersion in different sentiment periods. With a Wilcoxon Rank-Sum test, we find that analysts’ dispersion
is significantly higher when sentiment is low except for IBM. This is consistent with the time series dispersion measures about the real economy. Dispersions vary dramatically among stocks, which are high in financial and utility companies but low in retail and manufacture ones.

Figure 3: Analysts’ dispersion in different sentiment periods

![Analysts' dispersion in different sentiment periods](image)

*Notes:* The figure shows the distribution of analysts’ dispersion for each individual stock during high sentiment subsample, the whole sample and the low sentiment subsample. When the text is red (blue) it means the mean of analysts’ dispersion is significantly higher using The Wilcoxon Rank-Sum Test.

Figure 4 is a similar plot of the Idio-Sys Ratio. All stocks except three have larger gamma-beta ratio in low sentiment periods. This is mainly driven by a soaring idiosyncratic variance when sentiment is high.

4 High-frequency econometric procedures

We already have derived the elasticity as shown in Equation 2.2. However, the model is stylish and we need to estimate the elasticity from the data we have: one minute price and volume data. We first present how volume and volatility can be estimated, and then continue to propose the econometric model from which we can estimate the elasticity and the effect of the explanatory variables. Finally, bootstrapping algorithms are shown for statistical inferences. Many of the ideas follow from BLX, but we need to account for the panel data and subsample analysis in this section.
Figure 4: the $\gamma$-$\beta$ ratio in different sentiment periods

Notes: The figure shows the distribution of $\gamma$-$\beta$ ratio for each individual stock during high sentiment subsample, the whole sample and the low sentiment subsample. When the text is red (blue) it means the mean of $\gamma$-$\beta$ ratio is significantly higher using The Wilcoxon Rank-Sum Test.

4.1 Price and volume process

We fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Let $(P_{m,t})_{t \geq 0}$ denote the logarithmic price process of the market asset and $(P_{i,t})_{t \geq 0}$ be the price process of individual stock $i$. $P_{m,t}$ and $(P_{i,t})_{t \geq 0}$ follow from continuous-time finance literature (see, e.g., Merton (1992) and Duffie (2001)), and are assumed to be jump-diffusion processes of the form

\begin{align}
    dP_{m,t} &= b_{m,t}dt + \sigma_{m,t}dW_{m,t} + dJ_{m,t}, \\
    dP_{i,t} &= b_{i,t}dt + \beta_{i,t}\sigma_{m,t}dW_{m,t} + \sigma_{i,t}dW_{i,t} + dJ_{i,t}
\end{align}

(4.1) (4.2)

where $b$ is an instantaneous drift process, $\sigma$ is a stochastic spot volatility process, $W$ is a Brownian motion, and $J$ is a pure jump process. We assume that $W_{m,t}$ is perpendicular to $W_{i,t}$ and $W_{i,t}$ is perpendicular to $W_{j,t}$ if $i \neq j$. In this way we decompose the price process of individual stocks as the sum of the market component and the idiosyncratic component. We do not specify the structure of the drift process or the jump process because they will not enter the spot volatility.

Let $T$ denote the sample span and $\Delta_n$ denote the sampling interval of the high-frequency data. Assume the prices are sampled at discrete times \( \{l \Delta_n : 0 \leq l \leq [T/\Delta_n]\} \). The corresponding
high-frequency asset returns is denoted by \( r_l \equiv P_l \Delta_n - P_{(l-1)\Delta_n} \).

Let \( V_{l\Delta_n} \) denote the trading volume within the high-frequency interval \( [(l-1)\Delta_n, l\Delta_n] \). Following [Li and Xiu (2016)], we consider a general state-space model for both the market index and individual stocks

\[
V_{m,l\Delta_n} = \mathcal{V}(\zeta_{m,l\Delta_n}, \epsilon_{m,l\Delta_n}) \tag{4.3}
\]
\[
V_{i,l\Delta_n} = \mathcal{V}(\zeta_{i,l\Delta_n}, \epsilon_{i,l\Delta_n}) \tag{4.4}
\]

where as stated in their paper, \( \zeta \) is a latent state process, \( (\epsilon_{\Delta_n}) \) are i.i.d. transitory shocks with distribution \( F_\epsilon \), and \( \mathcal{V}(\cdot) \) is a general transform function. The latent state process \( \zeta \) captures time-varying conditioning information such as the intensity of order arrival and the shape of the order size distribution. Note that we are not imposing any restrictions on the structure of \( \zeta_{i,l\Delta_n} \) (resp. \( \epsilon_{i,l\Delta_n} \)) and it does not need to be perpendicular to \( \zeta_{m,l\Delta_n} \) (resp. \( \epsilon_{m,l\Delta_n} \)). We are not imposing restrictions on the form of \( F_\epsilon \) either, and the distributions can be different for the market index and individual stocks.

With this specification, the volume intensity process is calculated by

\[
m_t = \int \mathcal{V}(\zeta_t, \epsilon) F_\epsilon (d\epsilon) \tag{4.5}
\]

which is the instantaneous conditional mean process of \( V \). This process is used as the instantaneous empirical analogue to the expected volume in the theoretical models discussed in Section 2, both for the market and the individual stocks. Correspondingly, the spot volatility processes \( \sigma \) is used as the instantaneous analogue of the return standard deviation.

### 4.2 Volume and volatility jump

With the setup above, we need to get the volume jump and volatility jump from one minute return and one minute volume data. Let \( \tau \) denote the pre-scheduled announcement time. All announcements in our sample are pre-scheduled to be released on a certain minute, so we suppose \( \tau = (l(\tau) - 1) \Delta_n \) for a particular integer \( l(\tau) \)\(^5\) where \( \Delta_n = 1 \) minute is the sampling interval of our intraday data. Then the event window is defined as \( [(l(\tau) - 1) \Delta_n, l(\tau) \Delta_n] \).

Correspondingly, the \( k_n \)-minute periods immediate before (after) the event is defined as the pre-event (post-event) window. Denote stock \( i \)'s return and volume in the \( [(k-1)\Delta_n, k\Delta_n] \) window by \( r_{i,k} \) and \( V_{i,k\Delta_n} \), respectively. Let \( m_{t\tau-} \) (\( m_{t\tau} \)) denote the pre-event (post-event) volume intensity,

---

\(^5\)All announcements in our sample are indeed scheduled to be on the minute.
and let $\sigma_{i\tau}$ denote the spot volatility, then they can be estimated as follows:

$$
\hat{m}_{i\tau} = \frac{1}{k_n} \sum_{k=1}^{k_n} V_i(l(\tau)-k) \Delta_n,
\hat{m}_{i\tau} = \frac{1}{k_n} \sum_{k=1}^{k_n} V_i(l(\tau)+k) \Delta_n,
$$

$$(4.6)
$$

We denote the corresponding log volatility and log volume jumps by $\Delta \log(\sigma_{i\tau}) \equiv \log(\sigma_{i\tau}) - \log(\sigma_{i\tau-})$ and $\Delta \log(m_{i\tau}) \equiv \log(m_{i\tau}) - \log(m_{i\tau-})$, respectively. The reason we use log difference is that both volume and volatility have time variation which is not of interest in this paper. For example, there is an obvious upward trend in volume for many stocks due to expanding high frequency trading. Volumes and volatilities can also be affected by liquidity and transaction costs, which are hardly affected by news announcements, and are not closely related to our topic either. By substracting pre-event volume (volatility) from post-event volume (volatility), we only take account of the "abnormal" volume and volatility in order to focus on our attention only to news-triggered trading behaviors. We nonparametrically estimate the log volatility and log volume jumps as follows:

$$
\Delta \log(m_{i\tau}) = \log(\hat{m}_{i\tau}) - \log(\hat{m}_{i\tau-}),
\Delta \log(\sigma_{i\tau}) = \log(\hat{\sigma}_{i\tau}) - \log(\hat{\sigma}_{i\tau-})
$$

$$(4.7)
$$

Figure 5 demonstrates the positive correlation between log volume jumps and log volatility jumps for Dow Jones 30 components.

### 4.3 Fixed effect panel regression

Instead of a univariate linear model, now we model the volume-volatility relationship of each stock as follows:

$$
\Delta \log(m_{i\tau}) = c_i + a_0^T D_{\tau} + (b_0 + a_1^T D_{\tau} + b_1^T X_{i\tau}) \cdot \Delta \log(\sigma_{i\tau}),
$$

$$(4.8)
$$

where $\Delta \log(m_{i\tau})$ is the jump in stock $i$’s log volume intensity and $\Delta \log(\sigma_{i\tau})$ is the jump in its spot volatility. $D_{\tau}$ are dummy variables that represent the four categories of announcements rather than FOMC announcements when we combine all announcements. $X_{i\tau}$ is comprised of the different explanatory/control variables employed in the estimation. That means, though stocks differ in intercepts and slopes, they share the same specification in slopes: the coefficients of the explanatory/control variables are the same across stocks, though the values of these variables differ. One may also notice that different categories also share the same coefficients on explanatory
Notes: The figure shows the scatter plot of log volume jumps on log volatility jumps for 27 Dow Jones 30 components and the market ETF.

variable. The reason is, when we did the same analysis on each individual category, we found similar coefficients for these explanatory variable, and thus we combine these categories to increase sample size, considering that estimation errors of estimated volume and volatility is already high.

Following the notation in BLX, consider the group $A$ which is a set of a total of $M$ announcement times. We also implement DID to take account of intraday patterns, as shown in Equation 1.9, where $C(\tau)$ is a control group that occurs at the same time on different days and $N_C$ is the number of the control groups. In our regressions, $C(\tau)$ is the set of the same time as $\tau$ on the 22 days without announcements before the event day. That means, we use the average log volume (volatility) jump
at the event time of the previous 22 no-announcement days as control.

\[
\Delta \log (m_{i\tau}) \equiv \Delta \log (m_{i\tau}) - \frac{1}{N_C} \sum_{\tau' \in C(\tau)} \Delta \log (m_{i\tau'}),
\]

\[
\Delta \log (\sigma_{i\tau}) \equiv \Delta \log (\sigma_{i\tau}) - \frac{1}{N_C} \sum_{\tau' \in C(\tau)} \Delta \log (\sigma_{i\tau'}),
\]

(4.9)

Therefore, our question to investigate what affects volume-volatility elasticity is resolved by a fixed-effect panel regression as follows:

\[
\Delta \log (m_{i\tau}) = c_i + a_0^\top D_\tau + (b_0 + a_1^\top D_\tau + b_1^\top X_{i\tau}) \cdot \Delta \log (\sigma_{i\tau}) + \varepsilon_{i\tau}
\]

(4.10)

Note that this panel regression is different from the classic ones in that we do not assume the sample size goes to infinity. Instead, the size of the stock is relatively small, compared to the long time horizon. Therefore, fixed effect model does not decrease efficiency much, but ensures consistency. The asymptotic of the regression is not achieved by sample size going to infinity, but by the infinitely small estimation window around announcements.

Besides the panel regression, we also run regressions for the market ETF and treasury bond futures. These are the univariate version of the above equation:

\[
\Delta \log (m_{\tau}) = c + a_0^\top D_\tau + (b_0 + a_1^\top D_\tau + b_1^\top X_\tau) \cdot \Delta \log (\sigma_\tau) + \varepsilon_\tau
\]

(4.11)

For the market index, we also run regressions for FOMC announcements individually. In this case, \(D_\tau = 0\).

Our goal is therefore again to conduct valid inference about the parameter vector \(\theta \equiv (a_0, b_0, a_1, b_1)\), especially the components \(b_0\) and \(b_1\) that shows the benchmark elasticity and the effect of explanatory variables.

Moreover, we may estimate the parameter vector \(\theta \equiv (a_0, b_0, a_1, b_1)\) in (4.10) for the group \(A\) using the following least-square estimator

\[
\hat{\theta} \equiv \arg\min_{\theta} \sum_{i=1}^{n} \sum_{\tau \in A} \left( \phi \left( \Delta \log (m_{i\tau}) \right) - \phi \left( a_0^\top D_\tau + (b_0 + a_1^\top D_\tau + b_1^\top X_{i\tau}) \cdot \Delta \log (\sigma_{i\tau}) \right) \right)^2
\]

(4.12)

where

\[
\phi (Y_{i\tau}) = Y_{it} - \frac{1}{M} \sum_{\tau \in A} Y_{it}
\]

The estimator \(\hat{\theta}\) needs the information set of \(A(\tau)\) and \(C(\tau)\). Define \(S_\tau \equiv \cup_{i=1}^{n} (m_{i\tau -}, m_{i\tau}, \sigma_{i\tau -}, \sigma_{i\tau}, X_{i\tau})_{i \in \mathcal{N}}\). Then \(S \equiv (S_\tau)_{\tau \in A \cup C}\), contains all information needed for the estimation. Our estimator of \(S\) may then be expressed as \(\hat{S}_n \equiv (\hat{S}_\tau)_{\tau \in A \cup C}\), where \(\hat{S}_\tau = \cup_{i=1}^{n} (\hat{m}_{i\tau -}, \hat{m}_{i\tau}, \hat{\sigma}_{i\tau -}, \hat{\sigma}_{i\tau}, X_{i\tau})_{i \in \mathcal{N}}\). \(X_{i\tau}\) are the
explanatory variables as described in Section 3, and $\hat{m}_{i\tau-}, \hat{m}_{i\tau}, \hat{\sigma}_{i\tau-}, \hat{\sigma}_{i\tau}$ are the pre and post event volume intensity or spot volatility as defined in (4.6). Then this estimator can be expressed as $\widehat{\theta} = f(\hat{S})$.

In section 7, we show that $\hat{S}$ is a consistent estimator of $S$, which in turn implies that $f(\hat{S})$ consistently estimates $f(S)$, provided $f(\cdot)$ is a smooth function of the estimated quantities.

The bootstrap methods follow from BLX, except that we use vectors of returns and volumes to make sure that we keep the co-occurrences of tradings across different stocks. The validity of the bootstrap algorithms depends still on the locally i.i.d. assumption of the resampled vectors of volumes and returns, from the pre-announcement window and post-announcement window. The bootstrapped sample size is 1000 in all regressions.

**Bootstrap Algorithm**

Step 1: For each $\tau \in \mathcal{T}$, generate i.i.d. draws $(V_{i(\tau)-k}, r_{i(\tau)-j})_{1 \leq k \leq k_n}$ and $(V_{i(\tau)+k}, r_{i(\tau)+j})_{1 \leq k \leq k_n}$ from $(V_{i(\tau)-k}, r_{i(\tau)-j})_{1 \leq k \leq k_n}$ and $(V_{i(\tau)+k}, r_{i(\tau)+j})_{1 \leq k \leq k_n}$, respectively. $V_t$ is the vector of one minute volumes of the Dow Jones 30 components, while $r_t$ is the vector of one minute returns of the Dow Jones 30 components.

Step 2: Compute $\Delta \log(\hat{m}_{i\tau})^*$ and $\Delta \log(\hat{\sigma}_{i\tau})^*$ the same way as $\Delta \log(\hat{m}_{i\tau})$ and $\Delta \log(\hat{\sigma}_{i\tau})$, respectively, except that the original data $(V_{i(\tau)-k}, r_{i(\tau)-j})_{1 \leq k \leq k_n}$ is replaced with $(V_{i(\tau)-k}^*, r_{i(\tau)-j}^*)_{1 \leq k \leq k_n}$.

Rerun the regression according to (4.10) and get $b^*$. 

Step 3: Repeat steps 1 and 2 a large number of times. Report the empirical standard errors of each component of $b^* - b$ as the standard errors of the original estimator $b$.

**Differences of coefficients in high and low sentiment periods**

In this paper, we would like to investigate how regression results differ in high and low sentiment periods. Though the two samples do not overlap, they may share the same days as control, so the coefficients in the two subsamples are not independent. Therefore, we implement the following algorithm to generate statistical inference for the differences in coefficients in different sentiment periods:

Step 1: Separate the whole sample into high sentiment subsample and low sentiment one. Run regressions for each subsample and get coefficients $b_H$ and $b_L$. Define the difference as $b_{diff} = b_H - b_L$.

Step 2: In each round of bootstrap, we repeat the above and get $b_{diff}^{(s)}$, where $s = 1, 2, \cdots, 1000$. 

Step 3: Use the empirical standard errors of $b_{diff}^{(s)}$ as the standard errors of $b_{diff}$.

The validity of the bootstrap algorithms is proved in Section 7.
5 Volume-volatility relationship around public announcements

5.1 Jumps and announcements

In this section, we present the empirical results for the S&P500 ETF (SPY) which represents the market index, the 10-year treasury bond futures (TY) which represents the risk free assets, as well as the Dow Jones 30 components (DJ30) which are individual stocks. We first take a look at the log jumps of all these assets, and then study the regression results one by one.

5.2 Log volatility and volume jumps

Before we delve into the volume-volatility relationship in different sentiment periods, we check how the behavior of volume and volatility differs when sentiment is high and low. As can be seen from Figure 7, the difference between the two periods seem small for volatility. However, as shown in Figure 6, TY, SPY and more than one third of the stocks experience higher average volume jump during high sentiment periods. An interpretation for this is that investors’ dispersion is higher when sentiment is high. Garcia (2013) argues that in the literature of behavioral finance, volume jump should be zero when public information is released, if investors are fully rational and possess symmetric information. Therefore, one reason of the higher volume jump may be the large inflow of less rational investors when sentiment is high. Besides, one can also notice much higher volume and volatility jumps around FOMC announcements, highlighting its importance for not only S&P 500 ETF and treasury bond futures, but also the individual stocks.

5.3 Volume-volatility elasticities around public news announcements

The results for regression 4.11 are reported in Table 3. To save space we do not report the coefficients of the dummy variables, so $b_0$ is the elasticity of volume-volatility elasticity around FOMC announcements. From the top to the bottom panel are results for the whole sample, high sentiment periods, low sentiment periods and the difference between high and low sentiment periods. The sentiment periods are separated according to the Baker and Wurgler (2006) sentiment index, and the differences and their statistics are generated by the algorithm in Section 4.3.

The results in the first panel resemble those in BLX, where both the coefficients for unemployment dispersion and monthly index are significantly negative. When they are put together, the signs remain negative but the unemployment dispersion variable does not have significant explanatory power. NVIX seems to be the strongest among the three, directly driving the elasticity to around 1 and rendering the other two explanatory variables insignificant. One interpretation of
Figure 6: Average log volume jumps in different sentiment periods

Notes: The figure shows the mean and the 95% confidence interval of the daily volume jumps of the Dow Jones 30 components, S&P 500 ETF (SPY) and ten year treasury bond futures (TY) around FOMC announcements and all announcements. For each asset, the color of the ticker being red (blue) indicates that the mean is higher during high (low) sentiment period.

A glance at the second and third panels will reveal the difference between the two sentiment periods: all the explanatory power of the disagreement measures only becomes significant in the high sentiment period. When we compare the results column by column, we notice that when no explanatory variables are added into the elasticity, it is slightly smaller in high sentiment period with insignificant difference. When the disagreement measures are added, the magnitudes of the coefficients in the high sentiment periods are around twice as large as those in the whole sample, and the coefficient of unemployment dispersion becomes significant even when MonthlyIndex is added in. NVIX still seems to dominate, driving the unemployment dispersion insignificant, but the latter’s negative signs remain. On the contrary, none of the coefficients are significant in the

this result is that the news implied NIX measure seem to proxy difference of opinions best around all macroeconomic announcements. Considering that NVIX and Monthly Index are very similar inherently, in the following regressions we use only NVIX and put Monthly Index into robustness tests.
Figure 7: Average log volatility jumps in different sentiment periods

Notes: The figure shows the mean and the 95% confidence interval of the daily volatility jumps of the Dow Jones 30 components, S&P 500 ETF (SPY) and ten year treasury bond futures (TY) around FOMC announcements and all announcements. For each asset, the color of the ticker being red (blue) indicates that the mean is higher during high (low) sentiment period.

low sentiment subsample individually and the differences are significant for all these coefficients.

When we control for the disagreement measures, we see that during high sentiment periods the elasticity gradually increases to around 1, the theoretical value when there is no disagreement, which means that these explanatory variables captures the disagreement quite well in high sentiment periods. In the low sentiment periods, on the contrary, the elasticity fluctuates but does not increases towards 1. The dramatically increased adjusted $R^2$ during high sentiment regime adds further evidence to the effectiveness of the explanatory variables, contrasting the seldom changed $R^2$ in the low sentiment period. We also conducted F test and find the p value during high sentiment periods is always less than 0.002 while in low sentiment periods it is never less than 0.1. With all these evidences we safely conclude that the explanatory variables only matter during high sentiment periods.

Next we focus our attention on FOMC announcements and report the results in Table 4. Similar to BLX, both the coefficients of unemployment dispersion and the monthly index are significantly
Table 3: Volume-volatility elasticity estimates using SPY Data

### Whole Sample

**Baseline estimates:**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.696**</th>
<th>0.837**</th>
<th>0.897**</th>
<th>0.933**</th>
<th>1.011**</th>
<th>1.022**</th>
<th>1.009**</th>
<th>1.019**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.066)</td>
<td>(0.075)</td>
<td>(0.079)</td>
<td>(0.080)</td>
<td>(0.083)</td>
<td>(0.084)</td>
<td>(0.083)</td>
<td>(0.084)</td>
<td></td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

- **UNEMP Dispersion**
  - $-0.056^{**}$
  - (0.015)
- **MonthlyIndex**
  - $-0.065^{**}$
  - (0.015)
- **NVIX**
  - $-0.090^{**}$
  - (0.016)

**$R^2$**

<table>
<thead>
<tr>
<th></th>
<th>0.547</th>
<th>0.549</th>
<th>0.550</th>
<th>0.550</th>
<th>0.553</th>
<th>0.553</th>
<th>0.553</th>
<th>0.552</th>
</tr>
</thead>
</table>

### High Sentiment Periods

**Baseline estimates:**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.696**</th>
<th>0.891**</th>
<th>0.982**</th>
<th>1.092**</th>
<th>1.089**</th>
<th>1.115**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.095)</td>
<td>(0.105)</td>
<td>(0.108)</td>
<td>(0.116)</td>
<td>(0.114)</td>
<td>(0.118)</td>
<td></td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

- **UNEMP Dispersion**
  - $-0.115^{**}$
  - (0.035)
- **MonthlyIndex**
  - $-0.146^{**}$
  - (0.034)
- **NVIX**
  - $-0.160^{**}$
  - (0.031)

**$R^2$**

<table>
<thead>
<tr>
<th></th>
<th>0.551</th>
<th>0.556</th>
<th>0.558</th>
<th>0.560</th>
<th>0.562</th>
<th>0.561</th>
</tr>
</thead>
</table>

### Low Sentiment Periods

**Baseline estimates:**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.757**</th>
<th>0.723**</th>
<th>0.581**</th>
<th>0.586**</th>
<th>0.738**</th>
<th>0.723**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.093)</td>
<td>(0.112)</td>
<td>(0.136)</td>
<td>(0.140)</td>
<td>(0.147)</td>
<td>(0.151)</td>
<td></td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

- **UNEMP Dispersion**
  - 0.011
  - (0.018)
- **MonthlyIndex**
  - 0.045
  - (0.024)
- **NVIX**
  - 0.064
  - (0.025)

**$R^2$**

<table>
<thead>
<tr>
<th></th>
<th>0.558</th>
<th>0.557</th>
<th>0.559</th>
<th>0.558</th>
<th>0.557</th>
<th>0.556</th>
</tr>
</thead>
</table>

### Difference between High and Low Sentiment Periods

**Baseline estimates:**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>-0.061</th>
<th>0.169</th>
<th>0.402*</th>
<th>0.506**</th>
<th>0.350</th>
<th>0.392*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.135)</td>
<td>(0.155)</td>
<td>(0.174)</td>
<td>(0.183)</td>
<td>(0.190)</td>
<td>(0.194)</td>
<td></td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

- **UNEMP Dispersion**
  - $-0.125^{**}$
  - (0.040)
- **MonthlyIndex**
  - $-0.191^{**}$
  - (0.041)
- **NVIX**
  - $-0.165^{**}$
  - (0.040)

**$R^2$**

<table>
<thead>
<tr>
<th></th>
<th>-0.140**</th>
<th>-0.140**</th>
<th>-0.140**</th>
<th>-0.140**</th>
<th>-0.140**</th>
<th>-0.140**</th>
</tr>
</thead>
</table>

The table reports the results from the DID jump regression in equation (4.11) for the specification $\Delta \log(m_t) = c + a_0 D_t + (b_0 + a_1 D_t + b_1 X_{t}) \cdot \Delta \log(\sigma_t) + \varepsilon_t$ Bootstrapped standard errors (1000 repetitions) are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.

Negative. NVIX still plays an important role and dominates the two. FOMC sentiment is also significantly negative, even when it is put together with NVIX and unemployment dispersion.

The patterns of the coefficients when the sample is split into two are similar as those for all
Table 4: Volume-volatility elasticity estimates around FOMC Announcements using SPY Data

### Whole Sample

#### Baseline estimates:

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.696**</th>
<th>0.884**</th>
<th>1.040**</th>
<th>1.057**</th>
<th>0.828**</th>
<th>1.043**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SE)</td>
<td>(0.064)</td>
<td>(0.082)</td>
<td>(0.098)</td>
<td>(0.100)</td>
<td>(0.072)</td>
<td>(0.099)</td>
</tr>
</tbody>
</table>

#### Estimates for explanatory variables in elasticity ($b_1$):

<table>
<thead>
<tr>
<th>UNEMP Dispersion</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.074**</td>
<td>-0.098**</td>
<td>-0.098**</td>
<td>-0.080**</td>
<td>-0.088**</td>
<td>-0.052*</td>
</tr>
<tr>
<td>(0.019)</td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.023)</td>
<td>(0.019)</td>
<td>(0.022)</td>
</tr>
</tbody>
</table>

#### $R^2$

| Whole Sample | 0.317  | 0.340  | 0.358  | 0.356  | 0.348  | 0.359  |

### High Sentiment Periods

#### Baseline estimates:

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.696**</th>
<th>1.042**</th>
<th>1.214**</th>
<th>1.247**</th>
<th>0.838**</th>
<th>1.108**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SE)</td>
<td>(0.091)</td>
<td>(0.112)</td>
<td>(0.132)</td>
<td>(0.135)</td>
<td>(0.092)</td>
<td>(0.145)</td>
</tr>
</tbody>
</table>

#### Estimates for explanatory variables in elasticity ($b_1$):

<table>
<thead>
<tr>
<th>UNEMP Dispersion</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.203**</td>
<td>-0.212**</td>
<td>-0.212**</td>
<td>-0.135**</td>
<td>-0.149**</td>
<td>-0.110**</td>
</tr>
<tr>
<td>(0.044)</td>
<td>(0.042)</td>
<td>(0.043)</td>
<td>(0.047)</td>
<td>(0.028)</td>
<td>(0.041)</td>
</tr>
</tbody>
</table>

#### $R^2$

| High Sentiment Periods | 0.312  | 0.420  | 0.419  | 0.442  | 0.391  | 0.451  |

### Low Sentiment Periods

#### Baseline estimates:

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.757**</th>
<th>0.736**</th>
<th>0.824**</th>
<th>0.807**</th>
<th>0.763**</th>
<th>0.808**</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SE)</td>
<td>(0.093)</td>
<td>(0.126)</td>
<td>(0.183)</td>
<td>(0.188)</td>
<td>(0.114)</td>
<td>(0.197)</td>
</tr>
</tbody>
</table>

#### Estimates for explanatory variables in elasticity ($b_1$):

<table>
<thead>
<tr>
<th>UNEMP Dispersion</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.007</td>
<td>0.016</td>
<td>-0.016</td>
<td>-0.021</td>
<td>-0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.034)</td>
<td>(0.035)</td>
<td>(0.025)</td>
<td>(0.030)</td>
<td>(0.029)</td>
</tr>
</tbody>
</table>

#### $R^2$

| Low Sentiment Periods | 0.364  | 0.352  | 0.353  | 0.341  | 0.352  | 0.327  |

### Difference between High and Low Sentiment Periods

#### Baseline estimates:

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>-0.061</th>
<th>0.306</th>
<th>0.390</th>
<th>0.440</th>
<th>0.075</th>
<th>0.301</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SE)</td>
<td>(0.131)</td>
<td>(0.167)</td>
<td>(0.225)</td>
<td>(0.232)</td>
<td>(0.147)</td>
<td>(0.245)</td>
</tr>
</tbody>
</table>

#### Estimates for explanatory variables in elasticity ($b_1$):

<table>
<thead>
<tr>
<th>UNEMP Dispersion</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>UNEMP Dispersion</th>
<th>NVIX</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.209**</td>
<td>-0.196**</td>
<td>-0.196**</td>
<td>-0.114*</td>
<td>-0.146**</td>
<td>-0.109*</td>
</tr>
<tr>
<td>(0.051)</td>
<td>(0.054)</td>
<td>(0.056)</td>
<td>(0.054)</td>
<td>(0.040)</td>
<td>(0.050)</td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation (4.11) for the specification $\Delta \log(m_\tau) = c + (b_0 + b_1' X_\tau) \cdot \Delta \log(\sigma_\tau) + \epsilon_t$. Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
announcements. All the explanatory variables individually are significantly negative, with much larger magnitude in high sentiment periods. Unemployment dispersion is now significantly negative even when other variables are added in, highlighting the importance of the dispersion about real economy around FOMC announcements. FOMC sentiment’s effect dominates that of NVIX, which is consistent with BLX that FOMC sentiment, a tone measurement of the FOMC statements themselves, is more closely related to FOMC announcements. Therefore, for FOMC announcements the FOMC sentiment measures seem to be a better measure of uncertainty compared with NVIX, as the latter is a more general uncertainty measures.

A sharp contrast to these highly significant variables are the ones in the low sentiment period, none of which are significant anymore. Also contrasting to the increasing adjusted $R^2$ in the high sentiment period are the continuously dropping adjusted $R^2$ when more variables are added in. The F test shows that the p value is always less than 0.02 in the high sentiment periods and is never smaller than 0.4 during the low sentiment regime.

The last panel further confirms the argument that disagreement measures only matter during high sentiment period. This is consistent with our hypothesis that only when $h$ is high will disagreements significantly reduce elasticities. Intuitively, when sentiment is high, investors’ portfolios are performing well. Due to self-attribution they prone to be over-confident, which means $h$ is large and magnifies the relative disagreement term. On the contrary, low sentiment period usually corresponds to recessions, during when investors suffer from losses. Under such circumstances they become frustrated and regretful, and gradually lose confidence in themselves. $h$ is so small in this periods that relative disagreement will no longer cast effect in volume-volatility elasticity. Another possible explanation for this disparity is that, as argued in [Yu and Yuan (2011)], when sentiment is high there is higher proportion of sentiment traders, who tend to be inexperienced and naive, and thus more likely to be victims of overconfidence and self–attribution. When market is filled more of investors with high $h$, their belief dispersion plays a larger role around information arrival.

5.4 Dow Jones 30 Components

In this section, we study what will affect individual stocks’ volume-volatility elasticity. Individual stocks are more complicated than the market index in that there can be variations in information asymmetry and idiosyncratic risks over time and across stocks. In our model, the systematic disagreement term is $h \frac{|\Delta \mu|}{\sigma_{pm}}$. $|s_1 - s_2|$ is the difference in the prior precision, which can be viewed as asymmetry in private information. Note that $\frac{|\Delta \mu|}{\sigma_{pm}}$ is the difference in normalized
interpretations of the public information, and we define \( \frac{\Delta \mu}{\sigma_{pm(s_1-s_2)}} \) as the relative disagreement measure in Section \( \boxed{3} \) which is what really matters in this model.

For market index, asymmetry in priors may not play a large role over time, since it is hard to access private information about the market over time. Even if different investors possess asymmetric information in different stocks, they will aggregate out in the market portfolio. Therefore, it is reasonable to assume relatively stable information asymmetry along time. However, this is not the case for individual stocks. Investors may actively acquire private information, so there may be both time series variation and cross-sectional variation in information asymmetry for individual stocks, which we need to take account of.

To characterize this, we construct a ”relative disagreement” measure, defined as follows:

\[
\text{Relative } X_{i\tau} = X_\tau - AFD_{i\tau}
\]

\( AFD_{i\tau} \) is the analysts’ dispersion for stock \( i \) at announcement time \( \tau \). As described in Section \( \boxed{3.4} \) it is the monthly standard deviation of one year ahead forecasts for EPS, normalized by the average price of the same month. Therefore, \( AFD_{i\tau} \) is comparable across stocks and over time. \( X_\tau \) is the disagreement measure at time \( \tau \). To make \( AFD_{i\tau} \) comparable to \( X_\tau \) while keep its cross-sectional as well as time series variation, we devide each \( AFD_{i\tau} \) by a normalizing factor \( \sigma_{AFD} \), where

\[
\sigma_{AFD} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{T-1} \sum_{\tau=1}^{T} \left( AFD_{i\tau} - \frac{1}{T} \sum_{\tau=1}^{T} AFD_{i\tau} \right) \right)^2
\]

The process is that, we find the standard deviation of each time series, and then average across stocks to find the mean standard deviation \( \sigma_{AFD} \). After dividing each \( AFD_{i\tau} \) by \( \sigma_{AFD} \), \( \frac{1}{n} \sum_{i=1}^{n} \sigma_i = 1 \), where \( \sigma_i \) is the standard deviation of stock \( i \) over the whole sample. Recall that the standard deviation of \( X_\tau \) is 1, this normalization makes sure that \( X_\tau \) has similar magnitude of time variation as the average \( AFD_{i\tau} \). Then we get Relative \( X_{i\tau} \) and do the same normalization so that its time series standard deviation averaged across stocks is 1.

The reason for this subtraction is that analysts’ dispersion contain both disagreement and information asymmetry. The difference constructed above will be higher if asymmetric information is low, and lower when asymmetric information is high, and increases with the disagreement measure, which is quantitatively similar to the relative disagreement measure \( \frac{\Delta \mu}{\sigma_{pm(s_1-s_2)}} \). We avoid taking ratios because analysts’ dispersion can be 0 or close 0 for many stocks in a certain period, and the ratio will explode, driving the regression results invalid.
Table 5 displays the results for regression 4.10. The elasticity of individual stocks is much lower than that of the market, indicating higher dispersions of investors regarding the valuation of individual stocks. It also supports case 1 in model which predicts that investors who are optimistic about the systematic component also seem to be optimistic about the idiosyncratic part. NVIX is significantly negative, as predicted by the model. Its magnitude is much smaller than that of SPY, consistent with the fact that systematic term only composes a part of stocks’ pricing. Unemployment is negative but insignificant. Considering the multicollinearity by construction since analysts’ dispersion is subtracted from both of the disagreement measures, we do not put more than one disagreement measure in the model.

Next we turn to the Idio-Sys Ratio, and see a significantly negative sign. The magnitude is also much larger than that of the disagreement measures. The interpretation is that a similar unit of standard deviation in Idio-Sys Ratio will cast a much more negative effect than the disagreement measures. In our model, stocks with a high ratio of the idiosyncratic component relative to the systematic component will have lower elasticity. Intuitively, when the idiosyncratic part of a stock is high, it will be more difficult for investors to reach a consensus of the firms’ value due to this hard-to-estimate idiosyncratic component. We also notice that even when we have put in both the relative disagreement proxy and Idio-Sys Ratio, the elasticity is still far from 1, indicating that further improvement on this model may be needed to account for disagreements on pricing individual stocks.

The contrast between the coefficients of disagreement measures in panel 2 and panel 3 resembles that of SPY. When sentiment is high, all disagreement measures are significantly negative, with magnitudes several times of those for the whole period. On the contrary, the coefficient of unemployment dispersion is not significant in the low sentiment period, while the coefficients NVIX is much smaller in terms of absolute value. In high sentiment period both relative disagreement proxies increase adjusted $R^2$ while neither of them does so in low sentiment period. Also, elasticity increases by around 0.06 if we have controlled for the effect of one relative disagreement proxy while the value is less than 0.02 in low sentiment period. This disparity in results is consistent with the model implication: during low sentiment period the increased uncertainty lowers investors’ confidence in interpreting public signals, so the effect of disagreement measures wanes. Interestingly, this decrease in explanatory power is made up for by Idio-Sys Ratio, and both elasticities increase by around 0.1 when both the disagreement measures and Idio-Sys Ratio are added in. The result that Idio-Sys Ratio plays a larger role is consistent with the model implication that when confidence in interpreting the public signal is low, the idiosyncratic disagreement component is the only term
remained. All these differences are statistically significant, as manifested in panel 4.

5.5 Treasury Bond Futures

Finally we turn to the results for the 10 year treasury bond futures as reported in Table 6. Consider the effect of CPI on treasury bonds, we also include the one year ahead forecast dispersion on CPI. Over the whole sample, the regression results are very similar to those for SPY and individual stocks: all disagreement measures are significantly negative, and NVIX seems to play a dominating role again. The elasticity will be close to 1 when we have controlled for the disagreement measures.

However, when divided into two subsamples, the results are different from those of SPY and Dow Jones 30 components. Most regressors are significant in both regimes, with similar magnitudes. The only exception is unemployment dispersion, which turns insignificantly positive in the low sentiment period. This may imply that dispersion in CPI acts as a better disagreement proxy for treasury bond futures. Except for the unemployment dispersion, the differences are insignificant between the two regimes according to the last panel. In both periods, adjusted $R^2$ increases with add-ins of explanatory variables individually, and the p value of F test is smaller than 0.01 once NVIX is added into the regression, in both periods. Besides, we see large increases in elasticity when any of the proxies have been taken control of.

The most obvious differences between the two sentiment periods are the constant and elasticity. In high sentiment periods, the log volume jump absent of volatility jump is significantly higher. However, the elasticity is lower during high sentiment period, even after the disagreement proxies are already controlled. Therefore, though dispersion proxies play similar roles, there are still some differences that are not captured by the difference-of-opinion model around macroeconomic announcements when sentiment regime differs.

What we can conclude from Table 6 is that treasury bond futures do not demonstrate a similar disparity in different sentiment regimes. As a "risk free" asset, treasury bill should be the least affected by sentiment, so it does not display the distinct pattern across subsamples as SPY does. Actually, with the "flight to quality" and "flight to liquidity" stories, investors may not really lose confidence when they hold treasury bonds during low sentiment period. Besides, the sentiment index used in this paper is the first component of principal component analysis of five proxies in stock and mutual fund market, which may not accurately measure the sentiment in of the treasury bond market. Besides, there is even no well-defined sentiment measure in the treasury bond market considering the rarity of speculative activities in the treasury bonds. Therefore, in our paper, the
treasury bond futures may just act as a benchmark, which by comparison, highlights the sharp effect sentiment has on stock market.

6 Robustness Test

In this section, we test for robustness by using other explanatory variables and redefining sentiment periods. We rerun the regressions using RGDP growth rate dispersion and Monthly Index measure instead of Unemployment rate dispersion and NVIX. Also, the one year ahead CPI dispersion is replaced by the ten year ahead CPI dispersion. To save space, we only keep the results of interest: rows that manifest the differences between the high and low sentiment periods.

The results are shown in Table 7, Table 8, Table 9, and Table 10. We can still see the disparity between high and low sentiment when RGDP growth rate dispersion and MonthlyIndex are used, though the coefficient difference between high and low sentiment for RGDP growth rate dispersion becomes insignificant for the S&P 500 ETF.

To further make sure that the disparity between high and low sentiment period is not a coincidence from sample classification, we redivide and redo the analysis for the S&P 500 ETF and Dow Jones 30 components. Instead of defining the high and low sentiment periods based on sample median, we use the 25% and 75% percent quantiles. That means, a month is classified into high (low) sentiment period if the previous month’s sentiment is no less (no more) than the 75% (25%) sample quantile. The results are shown in Table 11, Table 12, and Table 13.

The patterns are exactly the same for the S&P 500 ETF, though some of the differences between the two sentiment periods become insignificant around FOMC announcements, which can be explained by the reduced sample size. Actually, when we use the 75% and 25% percent quantiles as thresholds, the disparity in Dow Jones 30 components becomes even more obvious. All disagreement measures are significantly negative during high sentiment periods while lose significance when sentiment level is low.

7 Conclusion

This paper proposes a new channel through which sentiment can affect information processing in the financial market. During high sentiment periods, investors are confident in their interpretations of the same public signal, while the confidence level is low in low sentiment periods. As a result, dispersion in interpretations only casts effect in the volume-volatility relationship during high sentiment periods, both for the market ETF and individual stocks. Econometric tools based
on high frequency data are provided in this paper to empirically estimate the effect of disagreement measures on volume-volatility elasticity in high and low sentiment periods.

To derive the volume-volatility relationship of individual stocks around systematic news release, we incorporate a CAPM structure with heterogeneous beliefs into the difference-of-opinion model. The model predicts that the elasticity decreases with investors’ disagreement on common public signals and the ratio of idiosyncratic volatility. Besides, during periods when investors’ confidence in interpreting the signals is low, the disagreement measures no longer have explanatory power because disagreements on interpretations may no longer enter the decision making process during transactions.

Our empirical results are consistent with the model predictions. Dispersion on real economy growth and uncertainty measures significantly reduces elasticity, so does the ratio of idiosyncratic volatility to systematic volatility. When we separate the sample into high and low sentiment periods based on an sentiment measure in the stock market, we find that the disagreement measures significantly reduces volume-volatility elasticity only in high sentiment periods, both for the S&P 500 ETF and Dow Jones 30 components. The coefficient for idiosyncratic ratio is negative in both periods, though. For treasury bill futures as benchmarks, there is no disparity between high and low sentiment periods. The robustness tests use new disagreement measures and new classification criterion for classifying sentiment periods, and the results remain the same.
Table 5: Volume-volatility elasticity estimates around using DJ30 data

<table>
<thead>
<tr>
<th>Whole Sample</th>
<th>Baseline estimates:</th>
<th>Estimates for explanatory variables in elasticity ($b_1$):</th>
</tr>
</thead>
</table>
| Elasticity ($b_0$) | 0.534** 0.541** 0.554** 0.599** 0.604** 0.622** | UNEMP Dispersion  
  -0.005  
  -0.010**  
  -0.026** |
|                  | (0.018) (0.018) (0.019) (0.021) (0.021) (0.022) | NVIX  
  -0.004  
  -0.011**  
  -0.027** |
|                  |                              | (0.003) (0.003) (0.004) (0.004) |
| Idio-Sys Ratio  | -0.026** -0.026** -0.027** | Idio-Sys Ratio  
  -0.011**  
  -0.012*  
  -0.015* |
|                  | (0.004) (0.004) (0.004) | (0.008) (0.008) (0.006) |
| $R^2$           | 0.299 0.299 0.299 0.300 0.300 0.301 | $R^2$  
  0.298  
  0.299  
  0.299  
  0.300  
  0.300  
  0.300 |

<table>
<thead>
<tr>
<th>High Sentiment Periods</th>
<th>Baseline estimates:</th>
<th>Estimates for explanatory variables in elasticity ($b_1$):</th>
</tr>
</thead>
</table>
| Elasticity ($b_0$) | 0.458** 0.511** 0.527** 0.502** 0.544** 0.568** | UNEMP Dispersion  
  -0.040**  
  -0.040**  
  -0.015** |
|                      | (0.022) (0.025) (0.026) (0.028) (0.030) (0.032) | NVIX  
  -0.037**  
  -0.040**  
  -0.012* |
|                      |                              | (0.009) (0.009) (0.006) (0.006) |
| Idio-Sys Ratio  | -0.012* -0.015* | Idio-Sys Ratio  
  -0.015**  
  -0.012*  
  -0.015* |
|                      | (0.006) (0.006) (0.006) | (0.008) (0.008) (0.006) |
| $R^2$           | 0.299 0.299 0.299 0.300 0.300 0.301 | $R^2$  
  0.299  
  0.299  
  0.299  
  0.300  
  0.300  
  0.300 |

<table>
<thead>
<tr>
<th>Low Sentiment Periods</th>
<th>Baseline estimates:</th>
<th>Estimates for explanatory variables in elasticity ($b_1$):</th>
</tr>
</thead>
</table>
| Elasticity ($b_0$) | 0.617** 0.615** 0.635** 0.684** 0.681** 0.703** | UNEMP Dispersion  
  0.002  
  -0.008**  
  -0.032** |
|                      | (0.027) (0.028) (0.028) (0.029) (0.029) (0.030) | NVIX  
  0.002  
  -0.009**  
  -0.032** |
|                      |                              | (0.003) (0.003) (0.005) (0.005) |
| Idio-Sys Ratio  | -0.032** -0.032** | Idio-Sys Ratio  
  -0.032**  
  -0.032**  
  -0.032** |
|                      | (0.005) (0.005) (0.005) | (0.003) (0.003) (0.005) |
| $R^2$           | 0.303 0.303 0.303 0.304 0.304 0.305 | $R^2$  
  0.303  
  0.303  
  0.303  
  0.304  
  0.304  
  0.305 |

<table>
<thead>
<tr>
<th>Difference between High and Low Sentiment Periods</th>
<th>Baseline estimates:</th>
<th>Estimates for explanatory variables in elasticity ($b_1$):</th>
</tr>
</thead>
</table>
| Elasticity ($b_0$) | -0.159** -0.103** -0.108** -0.181** -0.137** -0.135** | UNEMP Dispersion  
  -0.041**  
  -0.032**  
  0.017* |
|                      | (0.034) (0.036) (0.037) (0.039) (0.040) (0.042) | NVIX  
  -0.039**  
  -0.031**  
  0.020** |
|                      |                              | (0.009) (0.009) (0.007) (0.007) |
| Idio-Sys Ratio  | -0.031** | Idio-Sys Ratio  
  0.018*  
  0.018*  
  0.018* |
|                      | (0.009) (0.009) (0.007) | (0.009) (0.009) (0.007) |

The table reports the results from the DID jump regression in equation 4.10 for the specification $\Delta \log (m_{i\tau}) = c_i + a_i' D_{\tau} + (b_0 + a_1' D_{\tau} + b_1' X_i) \cdot \Delta \log (\sigma_{i\tau}) + \epsilon_{i\tau}$. 

Bootstrapped standard errors are reported in parentheses.

* and ** indicate significance at the 5% and 1% level, respectively.
Table 6: Volume-volatility elasticity estimates using TY Data

<table>
<thead>
<tr>
<th>Whole Sample</th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline estimates:</strong></td>
<td><strong>Baseline estimates:</strong></td>
<td><strong>Baseline estimates:</strong></td>
<td><strong>Baseline estimates:</strong></td>
</tr>
<tr>
<td>Constant (a₀)</td>
<td>1.240**</td>
<td>1.505**</td>
<td>0.823**</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.083)</td>
<td>(0.123)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Elasticity (b₀)</td>
<td>0.524**</td>
<td>0.330**</td>
<td>0.841**</td>
</tr>
<tr>
<td>(0.061)</td>
<td>(0.072)</td>
<td>(0.098)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity (b₁):</td>
<td>Estimates for explanatory variables in elasticity (b₁):</td>
<td>Estimates for explanatory variables in elasticity (b₁):</td>
<td>Estimates for explanatory variables in elasticity (b₁):</td>
</tr>
<tr>
<td>UNEMP Dispersion</td>
<td>-0.499**</td>
<td>-0.965**</td>
<td>0.052</td>
</tr>
<tr>
<td>(0.15)</td>
<td>(0.017)</td>
<td>(0.040)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>CPI Dispersion</td>
<td>-0.666**</td>
<td>-0.049*</td>
<td>-0.077**</td>
</tr>
<tr>
<td>(0.17)</td>
<td>(0.023)</td>
<td>(0.027)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>NVIX</td>
<td>-0.123**</td>
<td>-0.100**</td>
<td>-0.165**</td>
</tr>
<tr>
<td>(0.085)</td>
<td>(0.019)</td>
<td>(0.031)</td>
<td>(0.033)</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.640</td>
<td>0.656</td>
<td>0.631</td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation 4.11 for the specification \( \Delta \log (m_\tau) = c + a_0 D_\tau + (b_0 + a_1 D_\tau + b_1 X_\tau) \cdot \Delta \log (\sigma_\tau) + \varepsilon_t \). Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
### Table 7: Volume-volatility elasticity estimates using SPY Data

**High Sentiment Periods**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.696**</th>
<th>0.737**</th>
<th>0.982**</th>
<th>0.922**</th>
<th>1.089**</th>
<th>1.019**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.098)</td>
<td>(0.141)</td>
<td>(0.113)</td>
<td>(0.145)</td>
<td>(0.116)</td>
<td>(0.149)</td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

<table>
<thead>
<tr>
<th>RGDP Growth Dispersion</th>
<th>-0.012</th>
<th>0.021</th>
<th>0.024</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.034)</td>
<td>(0.033)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MonthlyIndex</th>
<th>-0.146**</th>
<th>-0.152**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.033)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NVIX</th>
<th>-0.160**</th>
<th>-0.166**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$R^2$</th>
<th>0.551</th>
<th>0.551</th>
<th>0.558</th>
<th>0.557</th>
<th>0.562</th>
<th>0.561</th>
</tr>
</thead>
</table>

**Low Sentiment Periods**

<table>
<thead>
<tr>
<th>Elasticity ($b_0$)</th>
<th>0.757**</th>
<th>0.860**</th>
<th>0.581**</th>
<th>0.701**</th>
<th>0.738**</th>
<th>0.830**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.132)</td>
<td>(0.134)</td>
<td>(0.147)</td>
<td>(0.151)</td>
<td>(0.163)</td>
</tr>
</tbody>
</table>

**Estimates for explanatory variables in elasticity ($b_1$):**

<table>
<thead>
<tr>
<th>RGDP Growth Dispersion</th>
<th>-0.023</th>
<th>-0.034</th>
<th>-0.024</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.021)</td>
<td>(0.021)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MonthlyIndex</th>
<th>0.045</th>
<th>0.053*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.025)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NVIX</th>
<th>0.004</th>
<th>0.008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.026)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$R^2$</th>
<th>0.558</th>
<th>0.558</th>
<th>0.559</th>
<th>0.559</th>
<th>0.557</th>
<th>0.557</th>
</tr>
</thead>
</table>

**Difference between High and Low Sentiment Periods**

<table>
<thead>
<tr>
<th>RGDP Growth Dispersion</th>
<th>0.011</th>
<th>0.054</th>
<th>0.047</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.039)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MonthlyIndex</th>
<th>-0.191**</th>
<th>-0.205**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.042)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NVIX</th>
<th>-0.165**</th>
<th>-0.174**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.039)</td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation 4.11 for the specification $\Delta \log (m_t) = c + a_0^{b_0} D_{s_t} + (b_0 + a_1^{b_1} D_{s_t} + b_1^{b_1} X_{s_t}) \cdot \Delta \log (\sigma_t) + \varepsilon_t$.

Bootstrapped standard errors are reported in parentheses.

* and ** indicate significance at the 5% and 1% level, respectively.
Table 8: Volume-volatility elasticity estimates around FOMC Announcements using SPY Data

<table>
<thead>
<tr>
<th>High Sentiment Periods</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity ( (b_0) )</td>
<td>0.696**</td>
<td>1.048**</td>
<td>1.156**</td>
<td>1.356**</td>
<td>0.838**</td>
<td>1.312**</td>
</tr>
<tr>
<td>(0.094)</td>
<td>(0.171)</td>
<td>(0.133)</td>
<td>(0.191)</td>
<td>(0.097)</td>
<td>(0.189)</td>
<td></td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity ( (b_1) ):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGDP Growth Dispersion</td>
<td>-0.101*</td>
<td>-0.069</td>
<td>-0.118*</td>
<td>-0.118*</td>
<td>-0.069</td>
<td>-0.118*</td>
</tr>
<tr>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>-0.234**</td>
<td>-0.213**</td>
<td>-0.030</td>
<td>-0.030</td>
<td>-0.030</td>
<td>-0.030</td>
</tr>
<tr>
<td>(0.055)</td>
<td>(0.054)</td>
<td>(0.077)</td>
<td>(0.077)</td>
<td>(0.077)</td>
<td>(0.077)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>-0.149**</td>
<td>-0.149**</td>
<td>-0.150**</td>
<td>-0.150**</td>
<td>-0.150**</td>
<td>-0.150**</td>
</tr>
<tr>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.041)</td>
<td>(0.041)</td>
<td>(0.041)</td>
<td>(0.041)</td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.312</td>
<td>0.330</td>
<td>0.378</td>
<td>0.380</td>
<td>0.391</td>
<td>0.416</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low Sentiment Periods</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity ( (b_0) )</td>
<td>0.757**</td>
<td>0.868**</td>
<td>0.595**</td>
<td>0.722**</td>
<td>0.763**</td>
<td>0.799**</td>
</tr>
<tr>
<td>(0.092)</td>
<td>(0.154)</td>
<td>(0.159)</td>
<td>(0.173)</td>
<td>(0.111)</td>
<td>(0.199)</td>
<td></td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity ( (b_1) ):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGDP Growth Dispersion</td>
<td>-0.024</td>
<td>-0.029</td>
<td>-0.054</td>
<td>-0.054</td>
<td>-0.054</td>
<td>-0.054</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.029)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>0.041</td>
<td>0.055</td>
<td>0.071*</td>
<td>0.071*</td>
<td>0.071*</td>
<td>0.071*</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.034)</td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.035)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>-0.003</td>
<td>-0.038</td>
<td>-0.003</td>
<td>-0.038</td>
<td>-0.038</td>
<td>-0.038</td>
</tr>
<tr>
<td>(0.029)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.364</td>
<td>0.355</td>
<td>0.359</td>
<td>0.354</td>
<td>0.352</td>
<td>0.347</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difference between High and Low Sentiment Periods</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates for explanatory variables in elasticity ( (b_1) ):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGDP Growth Dispersion</td>
<td>-0.076</td>
<td>-0.029</td>
<td>-0.064</td>
<td>-0.064</td>
<td>-0.064</td>
<td>-0.064</td>
</tr>
<tr>
<td>(0.052)</td>
<td>(0.052)</td>
<td>(0.057)</td>
<td>(0.057)</td>
<td>(0.057)</td>
<td>(0.057)</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>-0.276**</td>
<td>-0.268**</td>
<td>-0.101</td>
<td>-0.101</td>
<td>-0.101</td>
<td>-0.101</td>
</tr>
<tr>
<td>(0.063)</td>
<td>(0.064)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td>(0.085)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>-0.146**</td>
<td>-0.112*</td>
<td>-0.146**</td>
<td>-0.112*</td>
<td>-0.146**</td>
<td>-0.112*</td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.052)</td>
<td>(0.040)</td>
<td>(0.052)</td>
<td>(0.040)</td>
<td>(0.052)</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation 4.11 for the specification \( \Delta \log(m_t) = c + (b_0 + b_1 X_t) \cdot \Delta \log(\sigma_t) + \epsilon_t \). Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
Table 9: Volume-volatility elasticity estimates using DJ30 Data

<table>
<thead>
<tr>
<th></th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity ($b_0$)</td>
<td>0.458** 0.557** 0.509** 0.502** 0.583** 0.549**</td>
<td>0.617** 0.637** 0.633** 0.684** 0.703** 0.699**</td>
<td>-0.036** -0.033**</td>
</tr>
<tr>
<td></td>
<td>(0.024) (0.029) (0.027) (0.029) (0.033) (0.031)</td>
<td>(0.027) (0.027) (0.028) (0.029) (0.028)</td>
<td>(0.009) (0.009)</td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGDP Growth Dispersion</td>
<td>-0.045** -0.042**</td>
<td>-0.009** -0.009**</td>
<td>-0.036** -0.033**</td>
</tr>
<tr>
<td></td>
<td>(0.008) (0.008)</td>
<td>(0.003) (0.003)</td>
<td>(0.009) (0.009)</td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>-0.041**</td>
<td>-0.009**</td>
<td>-0.032** -0.032**</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.003)</td>
<td>(0.005) (0.005)</td>
</tr>
<tr>
<td>Idio-Sys Ratio</td>
<td>-0.015** -0.011 -0.014*</td>
<td>-0.032** -0.032** -0.032**</td>
<td>0.017* 0.021* 0.018*</td>
</tr>
<tr>
<td></td>
<td>(0.006) (0.006) (0.006)</td>
<td>(0.005) (0.005) (0.005)</td>
<td>(0.008) (0.008) (0.008)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.298 0.300 0.299 0.299 0.300 0.300</td>
<td>0.303 0.303 0.303 0.304 0.305 0.305</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation $\Delta \log(m_{i\tau}) = c_i + a_i D_x + (b_0 + a_i D_x + b_1 X_{i\tau}) \cdot \Delta \log(\sigma_{i\tau}) + \varepsilon_{i\tau}$ for the specification $\Delta \log(\sigma_{i\tau}) = c_i + a_i D_x + (b_0 + a_i D_x + b_1 X_{i\tau}) \cdot \Delta \log(\sigma_{i\tau}) + \varepsilon_{i\tau}$. Bootstrapped standard errors are reported in parentheses.

* and ** indicate significance at the 5% and 1% level, respectively.
Table 10: Volume-volatility elasticity estimates using TY Data

<table>
<thead>
<tr>
<th></th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity ($b_0$)</td>
<td>0.330** 0.474** 0.376** 0.748** 0.748**</td>
<td>0.841** 1.060** 1.010** 0.975** 1.061**</td>
<td>0.088* 0.128* 0.061 0.050 -0.050 0.128* 0.061 0.050 -0.050</td>
</tr>
<tr>
<td></td>
<td>(0.072) (0.100) (0.136) (0.141) (0.140)</td>
<td>(0.107) (0.118) (0.116) (0.117) (0.120)</td>
<td>(0.041) (0.034) (0.033) (0.033) (0.043)</td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGDP Growth Dispersion</td>
<td>-0.041* 0.001</td>
<td>-0.129** -0.126*</td>
<td>0.088*</td>
</tr>
<tr>
<td>CPI Dispersion</td>
<td>-0.010</td>
<td>-0.072**</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>-0.106**</td>
<td>-0.107**</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.656 0.656 0.655 0.660 0.659</td>
<td>0.631 0.636 0.634 0.633 0.635</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation 4.11 for the specification $\Delta \log (\tau_t) = \epsilon + a_0 D + (b_0 + a_1 X) \cdot \Delta \log (\sigma_t) + \epsilon_1$. Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
Table 11: Volume-volatility elasticity estimates using SPY Data (different sentiment subsamples)

<table>
<thead>
<tr>
<th></th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elasticity ($b_0$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.634**</td>
<td>0.687**</td>
<td>-0.053*</td>
</tr>
<tr>
<td></td>
<td>(0.143)</td>
<td>(0.148)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.966**</td>
<td>0.817**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.184)</td>
<td>(0.198)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.995**</td>
<td>0.672**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.173)</td>
<td>(0.239)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.260**</td>
<td>0.734**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.209)</td>
<td>(0.252)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.100**</td>
<td>0.686**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.186)</td>
<td>(0.224)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.212**</td>
<td>0.765**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.237)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Estimates for explanatory variables in elasticity ($b_1$):</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMP Dispersion</td>
<td>-0.129*</td>
<td>-0.031</td>
<td>-0.098</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>-0.160**</td>
<td>0.004</td>
<td>-0.164*</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.042)</td>
<td></td>
</tr>
<tr>
<td>NVIX</td>
<td>-0.147**</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.043)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.161**</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.130*</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.032)</td>
<td></td>
</tr>
<tr>
<td><strong>$R^2$</strong></td>
<td>0.467</td>
<td>0.479</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>0.473</td>
<td>0.476</td>
<td>0.479</td>
</tr>
<tr>
<td></td>
<td>0.476</td>
<td>0.479</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>0.477</td>
<td>0.476</td>
<td>0.476</td>
</tr>
<tr>
<td></td>
<td>0.477</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.477</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.476</td>
<td>0.476</td>
<td></td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Estimates for explanatory variables in elasticity ($b_1$):</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMP Dispersion</td>
<td>-0.031</td>
<td>-0.039</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>MonthlyIndex</td>
<td>0.004</td>
<td>0.028</td>
<td>-0.164*</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.043)</td>
<td></td>
</tr>
<tr>
<td>NVIX</td>
<td>0.028</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.146*</td>
<td>-0.161**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.056)</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation 4.11 for the specification $\Delta \log(m_\tau) = c + a_0 \Delta m_\tau + (b_0 + a_1 \text{D}_\tau + b_1 \text{X}_\tau) \cdot \Delta \log(\sigma_\tau) + \epsilon_\tau$. If the previous month’s sentiment level is above 75% (below 25%) of the sample sentiment level, this month is classified as high (low) sentiment period. Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
Table 12: Volume-volatility elasticity estimates around FOMC Announcements using SPY Data (different sentiment subsamples)

<table>
<thead>
<tr>
<th></th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity ($b_0$)</td>
<td>0.634** 1.152** 1.151** 1.260** 0.874** 1.241**</td>
<td>0.687** 0.883** 0.865** 0.954** 0.764** 0.946**</td>
<td>-0.155* -0.131 -0.134</td>
</tr>
<tr>
<td></td>
<td>(0.138) (0.209) (0.222) (0.242) (0.169) (0.280)</td>
<td>(0.146) (0.234) (0.252) (0.284) (0.171) (0.293)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMP Dispersion</td>
<td>-0.202** (0.063)</td>
<td>-0.047 (0.035)</td>
<td>-0.155* -0.131 -0.134</td>
</tr>
<tr>
<td>NVIX</td>
<td>-0.178** (0.066)</td>
<td>-0.037 (0.037)</td>
<td>-0.131 -0.134</td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>-0.064 (0.067)</td>
<td>-0.141 -0.134</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.043 -0.134</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.030 -0.134</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.078 (0.087)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.082 (0.086)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.178 (0.124)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.012 (0.239)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.209 0.290 0.244 0.270 0.234 0.242</td>
<td>0.208 0.191 0.185 0.158 0.185 0.121</td>
<td>0.185 0.158</td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation (4.11) for the specification $\Delta \log (m\_t) = c + (b_0 + b_1 \_X_t) \cdot \Delta \log (\sigma\_t) + \epsilon_t$. If the previous month’s sentiment level is above 75% (below 25%) of the sample sentiment level, this month is classified as high (low) sentiment period.

Bootstrapped standard errors are reported in parentheses.

* and ** indicate significance at the 5% and 1% level, respectively.
Table 13: Volume-volatility elasticity estimates around using DJ 30 Data (different sentiment subsamples)

<table>
<thead>
<tr>
<th></th>
<th>High Sentiment Periods</th>
<th>Low Sentiment Periods</th>
<th>Difference between High and Low Sentiment Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
<td>Estimates for explanatory variables in elasticity ($b_1$):</td>
</tr>
<tr>
<td>Elasticity ($b_0$)</td>
<td>0.442** 0.521** 0.515** 0.481** 0.538** 0.540**</td>
<td>0.615** 0.615** 0.619** 0.734** 0.733** 0.739**</td>
<td>-0.047** -0.046**</td>
</tr>
<tr>
<td></td>
<td>(0.029) (0.035) (0.037) (0.036) (0.040) (0.042)</td>
<td>(0.033) (0.033) (0.033) (0.039) (0.039) (0.039)</td>
<td>(0.012) (0.012)</td>
</tr>
<tr>
<td>UNEMP Dispersion</td>
<td>-0.048** -0.045**</td>
<td>-0.001 0.001</td>
<td>-0.038** -0.035**</td>
</tr>
<tr>
<td></td>
<td>(0.012) (0.012)</td>
<td>(0.003) (0.003)</td>
<td>(0.012) (0.012)</td>
</tr>
<tr>
<td>NVIX</td>
<td>-0.040** -0.038**</td>
<td>-0.002 0.003</td>
<td>-0.059** -0.059**</td>
</tr>
<tr>
<td></td>
<td>(0.012) (0.012)</td>
<td>(0.003) (0.003)</td>
<td>(0.009) (0.009)</td>
</tr>
<tr>
<td>Idio-Sys Ratio</td>
<td>-0.013 -0.007 -0.010</td>
<td>-0.059** -0.059**</td>
<td>0.046** 0.052**</td>
</tr>
<tr>
<td></td>
<td>(0.008) (0.008) (0.008)</td>
<td>(0.009) (0.009)</td>
<td>(0.012) (0.012)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.303 0.305 0.304 0.304 0.305 0.304</td>
<td>0.308 0.308 0.308 0.312 0.311 0.312</td>
<td>0.312 0.312 0.312</td>
</tr>
</tbody>
</table>

The table reports the results from the DID jump regression in equation (4.10) for the specification $\Delta \log (m_{i\tau}) = c_i + a_{i0} D_{\tau} + (b_0 + a_{i1} D_{\tau} + b_{i1} X_{i\tau}) \cdot \Delta \log (\sigma_{i\tau}) + \epsilon_{i\tau}$. If the previous month’s sentiment level is above 75% (below 25%) of the sample sentiment level, this month is classified as high (low) sentiment period. Bootstrapped standard errors are reported in parentheses. * and ** indicate significance at the 5% and 1% level, respectively.
Appendix A: Proofs for the econometric methods

This appendix presents the validity of the econometric methods and bootstraps proposed in Section 4.

A.1 Estimation and inference

We now begin to validate the estimation and inference. We focus on the most general estimation, \( \hat{\theta}_n \) defined by equation (4.10). We write \( \hat{\theta}_n = f(\tilde{S}_n) \), where \( f(\cdot) \) is defined implicitly by the definitions (4.9) and (4.10). Besides, \( S \equiv (S_T)_{T \in \mathcal{A}} \) contains all information needed for the estimation.

The following assumptions and notations are from BLX except the last one, which imposes finite bound on \( \beta_{it} \):

Assumption 2. (i) The price process \( P \) is given by (4.1) for \( J_t = \int_0^t \xi_s dN_s + \int_0^t \int_\mathbb{R} \delta(s, z) \mu(ds, dz) \), for both the market and individual stocks. The processes \( b_m, \sigma_m, b_i, \sigma_i \) are càdlàg (i.e., right continuous with left limit) and adapted; \( \sigma_{mt} \) and \( \sigma_{it} \) are positive for \( t \in [0, T] \) almost surely; the process \( \xi \) is predictable and locally bounded; \( N \) is a counting process that jumps at the scheduled announcement times which are specified by the set \( \mathcal{A} \); \( \delta \) is a predictable function; \( \mu \) is a Poisson random measure with compensator \( \nu(ds, dz) = ds \otimes \lambda(dz) \) for some finite measure \( \lambda \).

(ii) The volume process \( V \) satisfies (4.3). The process \( \zeta \) is càdlàg and adapted. The error terms \( (\epsilon_i) \) take values in some Polish space, are defined on an extension of \( (\Omega, \mathcal{F}) \), i.i.d. and independent of \( \mathcal{F} \).

(iii) For a sequence of stopping times \( (T_m)_{m \geq 1} \) increasing to infinity and constants \( (K_m)_{m \geq 1} \), we have \( E|\sigma_{t \wedge T_m} - \sigma_{s \wedge T_m}|^2 + E|\zeta_{t \wedge T_m} - \zeta_{s \wedge T_m}|^2 \leq K_m|t - s| \) for all \( t, s \) such that \( [s, t] \cap \mathcal{A} = \emptyset \).

(iv) \( \beta_{it} \) given by (4.2) satisfies that \( \sup_t |\beta_{it}| \leq K \).

Assumption 3 follows from BLX and is fairly standard in the study of high-frequency data. See BLX for detailed explanation. We add the assumption that the \( \beta \) process is bounded, which is not a binding consumption as empirically \( \beta \) is believed to be around 1. We also denote \( M_p(\cdot) \equiv \int \mathcal{V}(\cdot, \epsilon)^p F_\epsilon(d\epsilon) \) for \( p \geq 1 \) and denote \( v_{it} = M_2(\zeta_{it}) - M_1^2(\zeta_{it}) \). To accommodate to the multidimensional case, we also denote \( M_{ij}^2(\cdot, \cdot) = \int \mathcal{V}(\zeta_{it}, \epsilon_i)^p \mathcal{V}(\zeta_{jt}, \epsilon_j)^p F_{\epsilon_i, \epsilon_j}(d\epsilon_i, d\epsilon_j) \) and denote \( v_{t}^{i,j} = M_2(\zeta_{it}, \zeta_{jt}) - M_1(\zeta_{it}) M_1(\zeta_{jt}) \). \( F_{\epsilon_i, \epsilon_j} \) is the joint distribution of \( \epsilon_i \) and \( \epsilon_j \).

Assumption 4. (i) The function \( M_1(\cdot) \) is Lipschitz on compact sets and the functions \( M_2(\cdot), M_{ij}^2(\cdot, \cdot) \) and \( M_4(\cdot) \) are continuous.
Almost surely, the function $f$ is well-defined and continuously differentiable in a neighborhood of $\tilde{S}$.

Consider variables $(\eta_{i\tau}, \eta_{i\tau}, \eta'_{i\tau}, \eta'_{i\tau})_{\tau \in T}$ which, conditionally on $\mathcal{F}$, are mutually independent, centered Gaussian with variances $(\nu_{i\tau}, \nu_{i\tau}, \sigma^2_{i\tau} / 2, \sigma^2_{i\tau} / 2)_{i \in \mathcal{N}, \tau \in T}$, where $\mathcal{N}$ is the set of stock indexes. We denote the first differential of $f$ at $\tilde{S}$ with increment $d\tilde{S}$ by $F(\tilde{S}; \tilde{S}d)$. For a sequence $Y_n$ of random variables, we write $Y_n \overset{L_2}{\rightarrow} Y$ if $Y_n$ converges stably in law towards $Y$, meaning that $(Y_n, U)$ converges in distribution to $(Y, U)$ for any bounded $\mathcal{F}$-measurable random variable $U$.

Assumption 3) characterizes the growth rate of $k_n$. The undersmoothing condition eliminates the bias terms asymptotically, and thus facilitates the statistical inference. With these notations, we get a multi-dimensional version of the Theorem 1 in BLX as follows:

**Theorem 1.** (a) Under Assumptions 2, 3 and 4
\[
\sqrt{k_n}(\hat{\theta}_n - \theta) \overset{L_2}{\rightarrow} F(\hat{S}; (\eta_{i\tau}, \eta_{i\tau}, \eta'_{i\tau}, \eta'_{i\tau}, 0)_{i \in \mathcal{N}, \tau \in T}).
\] (A.1)

(b) Moreover, the conditional distribution function of $\sqrt{k_n}(\tilde{\theta}_n - \tilde{\theta})$ given the original data converges in probability to that of $F(\hat{S}; (\eta_{i\tau}, \eta_{i\tau}, \eta'_{i\tau}, \eta'_{i\tau}, 0)_{i \in \mathcal{N}, \tau \in T})$ under the uniform metric.

**A.2 Proofs**

**Proof of Theorem 1(a)**

The proof of directly follows from BLX. According to their Equation (A.5) and (A.8), for each stock $i$,
\[
\sqrt{k_n}(\hat{m}_{i\tau} - m_{i\tau}, \hat{m}_{i\tau} - m_{i\tau})_{\tau \in T} \overset{L_2}{\rightarrow} (\eta_{i\tau}, \eta_{i\tau})_{\tau \in T},
\] (A.2)
\[
\sqrt{k_n}(\hat{\sigma}_{i\tau} - \sigma_{i\tau}, \hat{\sigma}_{i\tau} - \sigma_{i\tau})_{\tau \in T} \overset{L_2}{\rightarrow} (\eta'_{i\tau}, \eta'_{i\tau})_{\tau \in T}.
\] (A.3)

By Proposition 5 of Barndorff-Nielsen et al. (2008) and the property of stable convergence in law, we can combine (A.2) and (A.3), yielding
\[
\sqrt{k_n}(\hat{S}_n - \tilde{S}) \overset{L_2}{\rightarrow} (\eta_{\tau}, \eta_{\tau}, \eta'_{\tau}, \eta'_{\tau}, 0)_{i \in \mathcal{N}, \tau \in T}.
\] (A.4)

The assertion (A.1) then follows from (A.4) and the delta method, since $f$ is still a smooth function.

**Q.E.D.**

**Proof of Theorem 1(b).** Step 1. We divide the proof into several steps. Denote $\mathcal{G} \equiv \mathcal{F} \vee
\[ \sigma \{ \epsilon_i : i \geq 0 \} \]. In this step, we show that, for each \( \tau \in \mathcal{T} \),

\[ \sqrt{k_n} \begin{pmatrix} m_{n,1}\tau - \hat{m}_{1}\tau \\ \vdots \\ m_{n,I}\tau - \hat{m}_{I}\tau \\ \sigma_{n,1}\tau - \hat{\sigma}_{1}\tau \\ \vdots \\ \sigma_{n,I}\tau - \hat{\sigma}_{I}\tau \end{pmatrix} \xrightarrow{\mathcal{G}} \begin{pmatrix} \eta_{1}\tau \\ \vdots \\ \eta_{I}\tau \end{pmatrix}. \] (A.5)

Notice that in this section we use \( I \) instead of \( n \) to denote the number of stocks, since \( n \) is used for asymptotic proofs. Observe that

\[ \sqrt{k_n} \begin{pmatrix} \hat{m}_{n,1}\tau - \hat{m}_{1}\tau \\ \hat{m}_{n,I}\tau - \hat{m}_{I}\tau \\ \hat{\sigma}_{n,1}\tau - \hat{\sigma}_{1}\tau \\ \vdots \\ \hat{\sigma}_{n,I}\tau - \hat{\sigma}_{I}\tau \end{pmatrix} = \frac{1}{k_n} \sum_{j=1}^{k_n} \begin{pmatrix} V_{1,\ell(l(\tau)-j)\Delta_n}^* \\ \vdots \\ V_{I,\ell(l(\tau)-j)\Delta_n}^* \\ r_{1,\ell(l(\tau)-j)/\Delta_n}^2 \\ \vdots \\ r_{I,\ell(l(\tau)-j)/\Delta_n}^2 \end{pmatrix} \xrightarrow{\mathcal{G}} \begin{pmatrix} V_{1,\ell(l(\tau)-j)\Delta_n} \\ \vdots \\ V_{I,\ell(l(\tau)-j)\Delta_n} \\ r_{1,\ell(l(\tau)-j)/\Delta_n}^2 \\ \vdots \\ r_{I,\ell(l(\tau)-j)/\Delta_n}^2 \end{pmatrix}. \] (A.6)

We bootstrap with replacement within a small window, so the summands in the right-hand side of (A.6) are i.i.d. with zero mean conditional on \( \mathcal{G} \) when the window size goes to infinitely small. We denote the \( \mathcal{G} \)-conditional covariance matrix of the left-hand side by

\[ \Sigma_{n,\tau} = \begin{pmatrix} H_{n,\tau} & Q_{n,\tau} \\ Q_{n,\tau}^T & Z_{n,\tau} \end{pmatrix} \]

For matrix \( X \), denote its \((i, j)\) element as \( X_{(i, j)} \), then

\[
\begin{align*}
H_{n,\tau}^{(ij)} &= \frac{1}{k_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} V_{j,\ell(l(\tau)-k)\Delta_n} - \left( \frac{1}{k_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} \right) \left( \frac{1}{k_n} \sum_{k=1}^{k_n} V_{j,\ell(l(\tau)-k)\Delta_n} \right), \\
Q_{n,\tau}^{(ij)} &= \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} r_{j,\ell(l(\tau)-k)\Delta_n}^2 - \left( \frac{1}{k_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} \right) \left( \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} r_{j,\ell(l(\tau)-k)}^2 \right), \\
Z_{n,\tau}^{(ij)} &= \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} r_{j,\ell(l(\tau)-k)}^2 - \left( \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} V_{i,\ell(l(\tau)-k)\Delta_n} \right) \left( \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} r_{j,\ell(l(\tau)-k)}^2 \right).
\end{align*}
\]

In step 2, we shall show that

\[
\begin{pmatrix} H_{n,\tau}^{(ij)} & Q_{n,\tau}^{(ij)} & Z_{n,\tau}^{(ij)} \end{pmatrix} \xrightarrow{\mathcal{G}} \begin{pmatrix} \nu_{\tau}^{(ij)} & 0 & 2\beta_{1,\tau}^2 \beta_{2,\tau}^2 \sigma_m^4 \end{pmatrix}. \] (A.7)

Then we get (A.5) following exactly the same arguments as in BLX.
Step 2. We aim to prove (A.7). Starting with $H_{n,r}^{(ij)}$, we do the following decomposition

$$\frac{1}{k_n} \sum_{k=1}^{k_n} V_{i,(l(r)-k)} \Delta_n V_{j,(l(r)-k)} \Delta_n = A_{1,n} + A_{2,n},$$

where

$$A_{1,n} = \frac{1}{k_n} \sum_{k=1}^{k_n} \left( V_{i,(l(r)-k)} \Delta_n V_{j,(l(r)-k)} \Delta_n - M_2(\zeta_{i,(l(r)-k)} \Delta_n, \zeta_{j,(l(r)-k)} \Delta_n) \right),$$

$$A_{2,n} = \frac{1}{k_n} \sum_{k=1}^{k_n} M_2(\zeta_{i,(l(r)-k)} \Delta_n, \zeta_{j,(l(r)-k)} \Delta_n).$$

Since each summand in $A_{1,n}$ is $\mathcal{F}$-conditionally independent with zero mean, by Hölder’s inequality we get

$$\mathbb{E} \left[ A_{1,n}^2 \big| \mathcal{F} \right] \leq K k_n^{-2} \sum_{k=1}^{k_n} \sqrt{M_2(\zeta_{i,(l(r)-k)} \Delta_n) M_2(\zeta_{j,(l(r)-k)} \Delta_n)} \leq K k_n^{-1} \rightarrow 0,$$

That means, $A_{1,n} = o_p(1)$. In addition, since $M_2(\cdot, \cdot)$ is continuous and $\zeta$ is càdlàg, $A_{2,n} \rightarrow M_2(\zeta_{i,r-}, \zeta_{j,r-})$. This together with (A.2) lead to

$$H_{n,r}^{(ij)} \xrightarrow{p} M_2(\zeta_{i,r-}, \zeta_{j,r-}) - M_1(\zeta_{i,r-}) M_1(\zeta_{j,r-}) = v_{ij}^{(r)}.$$

The proof for the limiting behavior of $Q_{n,r}^{(ij)}$ is very similar to that for $\Sigma_{n,r}^{(ij)}$ in BLX. We decompose

$$\frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} V_{i,(l(r)-k)} \Delta_n r_{j,(l(r)-k)}^2 \Delta_n = A_{3,n} + A_{4,n} + A_{5,n} + A_{6,n},$$

where

$$A_{3,n} = \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} \left( V_{i,(l(r)-k)} \Delta_n - M_1(\zeta_{i,(l(r)-k)} \Delta_n) \right) r_{j,l(r)-k}^2,$$

$$A_{4,n} = \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} \left( M_1(\zeta_{i,(l(r)-k)} \Delta_n) - M_1(\zeta_{i,(l(r)-k-1)} \Delta_n) \right) r_{j,l(r)-k}^2,$$

$$A_{5,n} = \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} M_1(\zeta_{i,(l(r)-k-1)} \Delta_n) \left( r_{j,l(r)-k}^2 - \mathbb{E} \left[ r_{j,l(r)-k}^2 \Delta_n \right] \right),$$

$$A_{6,n} = \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} M_1(\zeta_{i,(l(r)-k-1)} \Delta_n) \mathbb{E} \left[ r_{j,l(r)-k}^2 \Delta_n \right].$$

Then by replacing all $r_{i,l(r)-k}^2$ with $r_{j,l(r)-k}^2$ and all the proofs can easily follow. Then we safely conclude that

$$Q_{n,r}^{(ij)} \xrightarrow{p} 0.$$
Finally there comes the limiting behavior of $\mathbf{Z}_{n,\tau}^{(ij)}$. For each $\iota \in \{l(\tau) - k : 1 \leq k \leq k_n\}$, we can decompose $r_{i,\iota} = r'_{i,\iota} + r''_{i,\iota}$ in restriction to $\Omega_n$ where

$$dP_{m,t} = b_{m,t}dt + \sigma_{m,t}dW_{m,t} + dJ_{m,t}, \quad (A.8)$$

$$dP_{i,t} = b_{i,t}dt + \beta_{i,t}\sigma_{m,t}dW_{i,t} + \sigma_{i,t}dW_{i,t} + dJ_{i,t} \quad (A.9)$$

$$r'_{i,\iota} = \beta_{i,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n}(W_{m,(l-1)\Delta_n} - W_{m,(l-1)\Delta_n}) + \sigma_{i,(l-1)\Delta_n}(W_{i,(l-1)\Delta_n} - W_{i,(l-1)\Delta_n})$$

$$r''_{i,\iota} = \int_{(l-1)\Delta_n}^{\iota\Delta_n} b_{s}ds + \int_{(l-1)\Delta_n}^{\iota\Delta_n} (\beta_{i,s}\sigma_{m,s} - \beta_{i,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n})dW_{m,s} + \int_{(l-1)\Delta_n}^{\iota\Delta_n} (\sigma_{i,s} - \sigma_{(l-1)\Delta_n})dW_{i,s}.$$  

We can then decompose

$$\frac{1}{k_n\Delta_n^2}r_{i,l(\tau)-k}^2 = A_{7,n} + A_{8,n}, \quad (A.10)$$

where $A_{7,n} = \frac{1}{k_n\Delta_n^2} \sum_{k=1}^{k_n} r_{i,k(\tau)-k}^2$ and $A_{8,n}$ is defined implicitly by $[A.10]$. Here we assume $i \neq j$ as the case $i = j$ has already been proved in BLX. Note that

$$E[r_{j,i,t}^2r_{j,i,t}/\Delta_n^2|F(i-1)\Delta_n] = 3\beta_{i,(l-1)\Delta_n}^2\beta_{j,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n}^4 + \beta_{i,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n}^2\sigma_{j,(l-1)\Delta_n}^2 + \beta_{j,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n}^2\sigma_{j,(l-1)\Delta_n}^2.$$

The variance of the martingale difference sequence $r_{i,t}^2/r_{j,t}^2/\Delta_n^2 - E\left[r_{i,t}^2r_{j,t}^2/\Delta_n^2|F(i-1)\Delta_n\right]$ is bounded, $\beta_{it}$ is bounded, so it is then easy to show that

$$A_{7,n} \overset{p}{\rightarrow} 3\beta_{i,l-\tau}^2\beta_{j,l-\tau}\sigma_{m,l-\tau}^4 + \beta_{i,l-\tau}\sigma_{j,l-\tau}^2\sigma_{m,l-\tau}^2 + \beta_{j,l-\tau}\sigma_{i,l-\tau}^2\sigma_{m,l-\tau}^2 + \sigma_{i,l-\tau}\sigma_{j,l-\tau}^2.$$

Similarly, we have the following decomposition:

$$\frac{1}{k_n\Delta_n^2} \sum_{k=1}^{k_n} r_{i,l(\tau)-k}^2 = A_{9,n} + A_{10,n},$$

where $A_{9,n} = \frac{1}{k_n\Delta_n^2} \sum_{k=1}^{k_n} r_{i,k(\tau)-k}^2$. Then we can similarly prove that

$$A_{9,n} \overset{p}{\rightarrow} \beta_{i,l-\tau}\sigma_{m,l-\tau}^2 + \sigma_{i,l-\tau}^2.$$

Indeed, $E[r_{i,t}^2/\Delta_n|F(i-1)\Delta_n] = \beta_{i,(l-1)\Delta_n}\sigma_{m,(l-1)\Delta_n}^2 + \sigma_{i,(l-1)\Delta_n}^2.$

$$|A_{10,n}| \leq \frac{K}{k_n\Delta_n} \sum_{k=1}^{k_n} \left| r'_{i,l(\tau)-k} \right| \left| r''_{i,l(\tau)-k} \right| + \frac{K}{k_n\Delta_n} \sum_{k=1}^{k_n} \left| r''_{i,l(\tau)-k} \right|^2$$

By Hölder’s inequality,

$$E \left[ \left| r'_{i,l(\tau)-k} \right| \left| r''_{i,l(\tau)-k} \right| \right] \leq \sqrt{E \left[ \left| r'_{i,l(\tau)-k} \right|^2 \right] E \left[ \left| r''_{i,l(\tau)-k} \right|^2 \right]}$$

46
As \( \frac{1}{k_n \Delta_n} \sum_{k=1}^{k_n} r''_{i,l,(\tau)-k} = O_p(1) \), we only need to prove that \( \sum_{k=1}^{k_n} \left| r''_{i,l,(\tau)-k} \right|^2 = o_p(1) \). This can be concluded from \( \frac{1}{k_n \Delta_n^2} \sum_{k=1}^{k_n} r''_{i,l,(\tau)-k} = O_p(1) \), as shown in the appendix A of BLX.

With symmetry in \( i \) and \( j \),

\[
|A_{8,n}| \leq \frac{K}{k_n \Delta_n^2} \sum_{k=1}^{k_n} \left| r'_{i,l,(\tau)-k} \right|^2 \left| r'_{j,l,(\tau)-k} \right| \left| r''_{j,l,(\tau)-k} \right| + \frac{K}{k_n \Delta_n^2} \sum_{k=1}^{k_n} \left| r'_{i,l,(\tau)-k} \right| \left| r''_{i,l,(\tau)-k} \right| \left| r''_{j,l,(\tau)-k} \right|^2 + \frac{K}{k_n \Delta_n^2} \sum_{k=1}^{k_n} \left| r''_{i,l,(\tau)-k} \right|^2 \left| r''_{j,l,(\tau)-k} \right|^2.
\]

Apply Hölder’s inequality on each of the three components we get

\[
E \left[ \left| r'_{i,l,(\tau)-k} \right|^2 \left| r'_{j,l,(\tau)-k} \right| \left| r''_{j,l,(\tau)-k} \right| \right] \leq E \left[ \left| r'_{i,l,(\tau)-k} \right|^4 \right]^{1/2} E \left[ \left| r'_{j,l,(\tau)-k} \right|^4 \right]^{1/2} E \left[ \left| r''_{j,l,(\tau)-k} \right|^4 \right]^{1/4},
\]

\[
E \left[ \left| r'_{i,l,(\tau)-k} \right| \left| r''_{i,l,(\tau)-k} \right| \left| r''_{j,l,(\tau)-k} \right|^2 \right] \leq E \left[ \left| r''_{i,l,(\tau)-k} \right|^4 \right]^{1/4} E \left[ \left| r'_{j,l,(\tau)-k} \right|^4 \right]^{1/4} E \left[ \left| r''_{j,l,(\tau)-k} \right|^4 \right]^{1/2},
\]

\[
E \left[ \left| r''_{i,l,(\tau)-k} \right|^2 \left| r''_{j,l,(\tau)-k} \right|^2 \right] \leq \sqrt{E \left[ \left| r''_{i,l,(\tau)-k} \right|^4 \right] E \left[ \left| r''_{j,l,(\tau)-k} \right|^4 \right]}. 
\]

Therefore, we only need to prove that \( E \left[ \frac{1}{k_n \Delta_n^2} \sum_{j=1}^{k_n} \left| r''_{i,(\tau)-j} \right|^4 \right] \to 0 \). By the Burkholder–Davis–Gundy inequality and Hölder’s inequality,

\[
E \left[ \left| r''_{i,(\tau)-j} \right|^4 \right] \leq K \Delta_n^4 + \sup_t |\beta_t|^4 K \Delta_n E \left[ \int_{(l(\tau)-k-1) \Delta_n}^{(l(\tau)-k) \Delta_n} \left( \sigma_{m,s} - \sigma_{m,(l(\tau)-k-1) \Delta_n} \right)^4 ds \right] + K \Delta_n E \left[ \int_{(l(\tau)-k-1) \Delta_n}^{(l(\tau)-k) \Delta_n} \left( \sigma_{i,s} - \sigma_{i,(l(\tau)-k-1) \Delta_n} \right)^4 ds \right],
\]

\[
E \left[ \frac{1}{k_n \Delta_n^2} \sum_{j=1}^{k_n} \left| r''_{i,(\tau)-j} \right|^4 \right] \leq K \Delta_n^2 + K E \left[ \sup_{t \in [\tau-2k_n \Delta_n, \tau]} |\sigma_t - \sigma_{\tau-1}| \right] \to 0.
\]

Therefore, the third term of \((A.11)\) is \( o_p(1) \). As \( \frac{1}{k_n \Delta_n^2} \sum_{k=1}^{k_n} r''_{i,l,(\tau)-k} = O_p(1) \), we can conclude that the first two terms of \((A.11)\) are also \( o_p(1) \). Therefore, \( A_{8,n} = o_p(1) \). Hence,

\[
Z_{n,\tau-}^{(ij)} \overset{p}{\to} 3 \beta_2 \sigma_{m,\tau-} \sigma_{j,\tau-}^2 + \beta_2 \sigma_{m,\tau-}^2 \sigma_{j,\tau-} + \beta_2 \sigma_{m,\tau-} \sigma_{j,\tau-}^2 + \beta_2 \sigma_{i,\tau-}^2 \sigma_{m,\tau-} + \sigma_{i,\tau-}^2 \sigma_{j,\tau-}^2 \quad \text{and} \\
Z_{n,\tau-}^{(ij)} \overset{p}{\to} 2 \beta_2 \sigma_{m,\tau-} \sigma_{j,\tau-}^2 \sigma_{i,\tau-}^2.
\]

This marks the end of proof for \((A.7)\).
Step 3. For this step, we need to prove the multi-dimensional version of Equation (A.17) in BLX. Actually, by essentially the same argument for (A.5), we can show that

\[
\sqrt{k_n} \begin{pmatrix}
\hat{m}_{n,1\tau} - \hat{m}_{1\tau} \\
\cdots \\
\hat{m}_{n,I\tau} - \hat{m}_{I\tau} \\
\hat{\sigma}_{n,1\tau} - \hat{\sigma}_{1\tau} \\
\cdots \\
\hat{\sigma}_{n,I\tau} - \hat{\sigma}_{I\tau}
\end{pmatrix} \overset{\mathcal{L}\mathcal{G}}{\longrightarrow} \begin{pmatrix}
\eta_{1\tau} \\
\cdots \\
\eta_{I\tau} \\
\eta'_{1\tau} \\
\cdots \\
\eta'_{I\tau}
\end{pmatrix} \tag{A.14}
\]

The proof is finished.  \hspace{1cm} Q.E.D.
Appendix B: Derivation of the Model

B.1 The general case when there are n assets

We first derive the relationship between trading intensity and return volatility when there are n assets with a general variance distribution.

B.1.1 Prior to the release of the signal

Suppose that prior to the signal is released, investor \( j \) believes that the pricing of all stocks has the prior distribution as below:

\[
V_j \sim N(X_j, \Sigma_j),
\]

where \( V_j = \begin{bmatrix} V_{j1} & V_{j2} & \ldots & V_{jn} \end{bmatrix} \) is the vector of fundamental values. \( X_j = \begin{bmatrix} X_{j1} & X_{j2} & \ldots & X_{jn} \end{bmatrix} \) is a of means. \( \Sigma_j \) is the \( n \times n \) covariance matrix.

Each investor maximizes his expected utility

\[
\max_{m_i^{(1)}} \left[ -\exp \left\{ -\lambda m_i^{(1)'} (X_i - P^{(1)}) \right\} \right].
\]

The superscript \( (1) \) means the value corresponds to stage 1. \( P^{(1)} \) is the vector of prices before the investors see the signal. Then the optimal portfolio for investor \( j \) can be solved:

\[
m_j^{(1)} = \lambda^{-1} \Sigma_j^{-1} (X_j - P^{(1)})
\]

When market clears,

\[
\alpha m_1^{(1)} + (1 - \alpha) m_2^{(1)} = 0
\]

Then the equilibrium price \( P^{(1)*} \) and equilibrium positions \( m_j^{(1)*} \) can be solved.

B.1.2 The signal is released

Upon receiving the signal, for investor \( j \), the distribution of \( V \) interpreted from the signal \( L \) is \( N(\mu_j (L), \Omega_j) \). Then under a Bayesian framework, the posterior distribution is
\[ \mathbf{V}_j | \mathbf{L} \sim N \left( \mathbf{Y}_j, \left( \Sigma_j^{-1} + \Omega_j^{-1} \right)^{-1} \right) \]

\[ \mathbf{Y}_j = \rho_j \mathbf{X}_j + (\mathbf{I} - \rho_j) \mathbf{\mu}_j (L) \] where \[ \rho_j = \left( \Sigma_j^{-1} + \Omega_j^{-1} \right)^{-1} \Sigma_j^{-1} \]

We can similarly solve for the equilibrium prices and positions. If we use the following notations (\( L \) is suppressed for simplicity):

\[ \hat{\Sigma}^{-1} \equiv \alpha \Sigma_1^{-1} + (1 - \alpha) \Sigma_2^{-1} \]
\[ \hat{\Omega}^{-1} \equiv \alpha \Omega_1^{-1} + (1 - \alpha) \Omega_2^{-1} \]
\[ \hat{\mathbf{X}} \equiv \alpha \Sigma_1^{-1} \mathbf{X}_1 + (1 - \alpha) \Sigma_2^{-1} \mathbf{X}_2 \]
\[ \hat{\mathbf{\mu}} \equiv \alpha \Omega_1^{-1} \mathbf{\mu}_1 + (1 - \alpha) \Omega_2^{-1} \mathbf{\mu}_2 \]

The equilibrium prices and positions are

\[ \mathbf{P}^{(1)*} = \hat{\Sigma} \hat{\mathbf{X}} \]
\[ \mathbf{P}^{(2)*} = \left( \hat{\Sigma}^{-1} + \hat{\Omega}^{-1} \right)^{-1} \left( \hat{\mathbf{X}} + \hat{\mathbf{\mu}} \right) \]

The equilibrium holdings are

\[ \mathbf{m}_2^{(1)*} = \Sigma_1^{-1} \left( \mathbf{X}_1 - \mathbf{P}^{(1)*} \right) \]
\[ \mathbf{m}_1^{(2)*} = \lambda^{-1} \left( \Sigma_1^{-1} + \Omega_1^{-1} \right) \left( \mathbf{Y}_1 - \mathbf{P}^{(2)*} \right) \]

The price change is

\[ \Delta \mathbf{P}^* = \mathbf{P}^{(2)*} - \mathbf{P}^{(1)*} \]
\[ = \left( \hat{\Sigma}^{-1} + \hat{\Omega}^{-1} \right)^{-1} \left( \hat{\mathbf{\mu}} - \hat{\Omega}^{-1} \mathbf{P}^{(1)*} \right) \]  \hspace{1cm} (B.1)

Volume is the change in position:

\[ \text{Vol}^* = \alpha \left| \mathbf{m}_2^{(2)*} - \mathbf{m}_1^{(1)*} \right| \]
\[ = \lambda^{-1} \alpha (1 - \alpha) \left| \Omega_1^{-1} \hat{\Omega} \Omega_2^{-1} (\mathbf{\mu}_1 - \mathbf{\mu}_2) + (\Omega_2 \Sigma_2^{-1} - \Omega_1 \Sigma_1^{-1}) \Delta \mathbf{P}^* \right| \]  \hspace{1cm} (B.2)
Equation B.3 is obtained by plugging $P^{(1)^*}$ in as a function of $\Delta P^*$ from equation B.1.

Therefore, the equilibrium volume and price change relationship is as follows:

$$\text{Vol}^* = \lambda^{-1}\alpha (1 - \alpha) \left| \Omega^{-1}_1 \hat{\Omega}_2^{-1} (\mu_1 - \mu_2) + \Omega^{-1}_1 \hat{\Omega}_2^{-1} \left( \Omega_2 \Sigma_2^{-1} - \Omega_1 \Sigma_1^{-1} \right) \Delta P^* \right|.$$

### B.2 Implement the CAPM Model with heterogeneous beliefs

Suppose the pricing of all stocks for investor $j$ follows the distribution below:

$$V \sim N(X_j, \Sigma_j).$$

Hence after we incorporate the one factor structure with heterogeneous beliefs, the expressions of the parameters are:

$$X_j = \beta x_{mj} + \gamma \eta_j,$$

$$\Sigma_j = (\beta \beta' + S) z_{mj}^2,$$

where $\beta = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{bmatrix}$, $\gamma = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{bmatrix}$ and $S = \text{diag}(\begin{bmatrix} \gamma_1^2 & \gamma_2^2 & \cdots & \gamma_n^2 \end{bmatrix})$. As the structure of the covariance matrix persists, we can simplify the proof if we introduce the following matrix:

$$M = \begin{bmatrix} \beta_1^2 + \gamma_1^2 & \beta_1 \beta_2 & \cdots & \beta_1 \beta_n \\ \beta_2 \beta_1 & \beta_2^2 + \gamma_2^2 & \cdots & \beta_2 \beta_n \\ \vdots & \vdots & \ddots & \vdots \\ \beta_n \beta_1 & \beta_n \beta_2 & \cdots & \beta_n^2 + \gamma_n^2 \end{bmatrix}.$$

It can be regarded as the covariance coefficients. The joint distribution of the stock prices can be written as functions of $\beta, \gamma, M$ as well as the parameters of the distribution of the market payoff.

$$X_j = \beta x_{mj} + \gamma \eta_j,$$

$$\Sigma_j = \sigma_{mj}^2 M,$$

$$\mu_j (L) = \beta (L - \mu_{mj}) + \gamma \eta_j,$$

$$\Omega_j = \sigma_{mj}^2 M.$$

Plug them in equation B.3 we get
Vol* = c \left| M^{-1} (\mu_1 - \mu_2) + \left( \frac{\sigma_{m_1}^2}{\sigma_{m_1}^2} - \frac{\sigma_{m_2}^2}{\sigma_{m_2}^2} \right) M^{-1} \Delta P^* \right|, \quad (B.4)

where \( c = \lambda^{-1} \alpha (1 - \alpha) (\alpha \sigma_{m_1}^{-2} + (1 - \alpha) \sigma_{m_2}^{-2})^{-1} \sigma_{m_1}^2 \sigma_{m_2}^2. \)

With the notation \( \Delta \mu_m = \mu_{m_2} - \mu_{m_1} \) and \( \Delta \eta = \eta_1 - \eta_2, \)

\[
\mu_1 (L) - \mu_2 (L) = \beta \Delta \mu_m + \gamma \Delta \eta.
\]

For any variable \( x, \) we define

\[
\begin{align*}
\tilde{x} &= \frac{\alpha \sigma_{m_1}^{-2} x_1 + (1 - \alpha) \sigma_{m_2}^{-2} x_2}{\alpha \sigma_{m_1}^{-2} + (1 - \alpha) \sigma_{m_2}^{-2}}, \\
\hat{x} &= \frac{\alpha z_{m_1}^{-2} x_1 + (1 - \alpha) z_{m_2}^{-2} x_2}{\alpha z_{m_1}^{-2} + (1 - \alpha) z_{m_2}^{-2}}, \\
\kappa_m &= \frac{\alpha \sigma_{m_1}^{-2} + (1 - \alpha) \sigma_{m_2}^{-2}}{\alpha \sigma_{m_1}^{-2} + (1 - \alpha) \sigma_{m_2}^{-2} + \alpha z_{m_1}^{-2} + (1 - \alpha) z_{m_2}^{-2}},
\end{align*}
\]

then the price change of the stocks is

\[
\Delta P^* = P^{(2)*} - P^{(1)*} = \left( \tilde{\Sigma}^{-1} + \hat{\Omega}^{-1} \right)^{-1} \left( \tilde{\mu} (L) - \hat{\Omega}^{-1} \tilde{\Sigma} \hat{x} \right) = \beta F_1 + \gamma F_2,
\]

where

\[
\begin{align*}
F_1 &= \kappa_m (L - \tilde{\mu} - \tilde{x}), \\
F_2 &= \kappa_m (\hat{\eta} - \tilde{\eta}).
\end{align*}
\]

We notice that \( F_1 = \Delta P_m^* \) has exact the same expression as in Kandel and Pearson (1995), so it can be treated as the equilibrium price change for the market.

\( F_2 \) is a constant, the value of which is

\[
\mu_\eta = \alpha (1 - \alpha) \left( \alpha + (1 - \alpha) \frac{\sigma_{m_1}^2}{\sigma_{m_2}^2} \right)^{-1} \left( \alpha \frac{z_{m_1}^2}{z_{m_2}^2} + (1 - \alpha) \right)^{-1} \left( 1 - \frac{\sigma_{m_1}^2}{\sigma_{m_2}^2} \frac{z_{m_1}^2}{z_{m_2}^2} \right) \Delta \eta.
\]

When the prices of stocks have such a factor structure, the distribution of their equilibrium prices is
\[ \Delta P^* \sim N(\gamma \mu_p, \Sigma_p), \]

where

\[ \Sigma_p = \begin{bmatrix}
\beta_1^2 & \beta_1 \beta_2 & \cdots & \beta_1 \beta_n \\
\beta_2 \beta_1 & \beta_2^2 & \cdots & \beta_2 \beta_n \\
\vdots & \vdots & \ddots & \vdots \\
\beta_n \beta_1 & \beta_n \beta_2 & \cdots & \beta_n^2
\end{bmatrix}. \]

### B.3 Volume-volatility elasticity

We aim to find the individual stock’s volume-volatility elasticity from equation B.4:

\[ \text{Vol}^* = c \left| M^{-1} (\mu_1 - \mu_2) + \left( \frac{\sigma^2_{m1}}{\sigma^2_{m1}} - \frac{\sigma^2_{m2}}{\sigma^2_{m2}} \right) M^{-1} \Delta P^* \right|. \]

According to Kan and Robotti (2017), if \( X \sim N(\mu, \Sigma) \) and \( Y = c |X| \),

\[ E(\text{Vol}^*_i) = \mu_i \text{erf} \left( \frac{1}{\sqrt{2} \sigma_i} \right) + 2\sigma_i \phi \left( \frac{\mu_i}{\sigma_i} \right) \]

\[ = \mu_i \left[ 2\Phi \left( \frac{\mu_i}{\sigma_i} \right) - 1 \right] + \sqrt{\frac{2}{\pi}} \sigma_i \exp \left\{ -\frac{\mu_i^2}{2\sigma_i^2} \right\}. \]

where \( \mu_i \) is the \( i \)th element of \( \mu \) and \( \sigma^2_i \) is the \( i \)th element on the diagonal of \( \Sigma \), so the volume-volatility elasticity is exactly the same as the univariate case in BLX. Since we are not interested in a particular equilibrium volume and price change, but instead in the "average" level, we continue to find out the relationship between expected volume and price change volatility.

Denote \( v_i = E[\text{Vol}^*_i] \), and \( \sigma_{p_i} \) be the volatility of \( p_i \), the price change of stock \( i \), then the relationship between \( v_i \) and \( \sigma_{p_i} \) is again

\[ \mathcal{E}_i \equiv \frac{\partial v_i(\sigma_{p_i})/v_i(\sigma_{p_i})}{\partial \sigma_{p_i}/\sigma_{p_i}} = \frac{1}{1 + \psi(\theta_i)}, \]

where \( \theta_i \) is the ratio of the mean to the volatility of the \( i \)th stock. Therefore, we need to find the distribution of \( \text{Vol}^* \) and \( \Delta P^* \). Let

\[ A \equiv \left( \frac{\sigma^2_{m1}}{\sigma^2_{m1}} - \frac{\sigma^2_{m2}}{\sigma^2_{m2}} \right) M^{-1}, B \equiv M^{-1} (\mu_1 - \mu_2). \]
then $\text{Vol}^* = c|B + A\Delta P^*|$ and the distribution of $B + A\Delta P^*$ is $N\left( A\gamma\mu + B, A\Sigma_p A^T \right)$. Without loss of generality, we look at the first stock. For the first stock, $\theta_1$ is the ratio of the first element of $A\gamma\mu + B$ to the square root of the $(1,1)$ element of $A\Sigma_p A^T$. We first look at $A\Sigma_p A^T$.

It is easy to prove that

$$M^{-1} = \left( \sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2 \right)^{-1} H,$$

where the $(k,k)$ element of $H$ is $\sum_{i \neq k} \beta_i^2 \gamma_j^2 + \prod_{i \neq k} \gamma_i^2$ and the $(k,l)$ ($k \neq j$) element of $H$ is $-\beta_1 \beta_2 \prod_{j \neq k,l} \gamma_j^2$.

Hence the $(1,1)$ element of $A\Sigma_p A^T$ is

$$\sqrt{f_{11}(\Sigma_p)} = \frac{\frac{\sigma^2_{m1}}{\sigma^2_{m1}} - \frac{\sigma^2_{m2}}{\sigma^2_{m2}}}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} \beta_1 \prod_{i=1}^{n} \gamma_i^2 \sigma_{p,m}.$$

Next we turn to the mean vector

$$B + A\gamma\mu = M^{-1}(\beta\Delta\mu + \gamma\Delta\eta) + \frac{\frac{\sigma^2_{m1}}{\sigma^2_{m1}} - \frac{\sigma^2_{m2}}{\sigma^2_{m2}}}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} M^{-1} \gamma\mu,$$

the first component of which is

$$\frac{\prod_{i \neq 1} \gamma_i^2 \beta_1}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} \Delta\mu + \left( \sum_{i \neq 1} \prod_{j \neq i,1} \beta_i^2 \gamma_j^2 + \prod_{i \neq 1} \gamma_i^2 \right) \frac{\gamma_1 - \beta_1 \sum_{i \neq 1} \prod_{j \neq i,1} \beta_i \gamma_i \gamma_j^2}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} \Delta\eta$$

$$+ \frac{\frac{\sigma^2_{m1}}{\sigma^2_{m1}} - \frac{\sigma^2_{m2}}{\sigma^2_{m2}}}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} \left( \sum_{i \neq 1} \prod_{j \neq i,1} \beta_i^2 \gamma_j^2 + \prod_{i \neq 1} \gamma_i^2 \right) \frac{\gamma_1 - \beta_1 \sum_{i \neq 1} \prod_{j \neq i,1} \beta_i \gamma_i \gamma_j^2}{\sum_{i=1}^{n} \beta_i \prod_{j \neq i} \gamma_j^2 + \prod_{i=1}^{n} \gamma_i^2} \mu.$$

Devide it by $\sqrt{f_{11}(\Sigma_p)}$ we get
\[ \theta_1 = \frac{\sigma^2_{m1}}{z^2_{m1}} - \frac{\sigma^2_{m2}}{z^2_{m2}} \frac{\Delta \mu}{\sigma_{pm}} + \frac{\gamma_1}{\beta_1} \left( \frac{\sum_{i=2}^{n} \beta_i^2}{\gamma_i^2} + 1 \right) \frac{\Delta \eta}{\sigma_{pm}} \\
+ \operatorname{sgn} \left( \frac{\sigma^2_{m1}}{z^2_{m1}} - \frac{\sigma^2_{m2}}{z^2_{m2}} \right) \left( \frac{\gamma_1}{\beta_1} \left( \frac{\sum_{i=2}^{n} \beta_i^2}{\gamma_i^2} + 1 \right) - \sum_{i=2}^{n} \frac{\beta_i}{\gamma_i} \right) \frac{\mu_{\eta}}{\sigma_{pm}}. \]

Recall that \( \mu_{\eta} = \alpha (1 - \alpha) \left( \alpha \sigma^2_{m2} + (1 - \alpha) \sigma^2_{m1} \right)^{-1} \left( \alpha z_{m1}^2 + (1 - \alpha) z_{m2}^2 \right)^{-1} \left( \frac{\sigma^2_{m2}}{z^2_{m2}} - \frac{\sigma^2_{m1}}{z^2_{m1}} \right) \Delta \eta. \) If we denote

\[ \Delta w = \left| \frac{\sigma^2_{m1}}{z^2_{m1}} - \frac{\sigma^2_{m2}}{z^2_{m2}} \right|, \]

\[ g \left( \frac{\gamma_i}{\beta_i} \right) = \frac{\gamma_i}{\beta_i} \left( \sum_{k=1}^{n} \frac{\beta_k^2}{\gamma_k^2} + 1 \right) - \sum_{j=1}^{n} \frac{\beta_k}{\gamma_k}, \]

\[ c_m = \alpha (1 - \alpha) \left( \alpha \sigma^2_{m2} + (1 - \alpha) \sigma^2_{m1} \right)^{-1} \left( \alpha z_{m1}^2 + (1 - \alpha) z_{m2}^2 \right)^{-1}, \]

then the elasticity of stock \( i \) can be written as

\[ \mathcal{E}_i = \frac{1}{1 + \psi(\theta_i)}; \]

\[ \theta_i = \Delta w^{-1} \frac{\Delta \mu}{\sigma_{pm}} + \left( \Delta w^{-1} + c_m \Delta w \right) g \left( \frac{\gamma_i}{\beta_i} \right) \frac{\Delta \eta}{\sigma_{pm}}. \]

To further simplify the analysis, consider a special case: when \( \sigma^2_{m1} = \sigma^2_{m2} = h^{-1} \), and \( s_j = z_{mj}^{-2} \) be the precisions, then

\[ \theta_i = \frac{h}{|s_1 - s_2|} \frac{\Delta \mu_m}{\sigma_{pm}} + \left( \frac{h}{|s_1 - s_2|} + \alpha (1 - \alpha) \left( \frac{|s_1 - s_2|}{\alpha s_1 + (1 - \alpha) s_2} \right) \right) g \left( \frac{\gamma_i}{\beta_i} \right) \frac{\Delta \eta}{\sigma_{pm}}. \]

Let

\[ \xi(h, s_1, s_2) \equiv \alpha (1 - \alpha) \left( \frac{|s_1 - s_2|}{\alpha s_1 + (1 - \alpha) s_2} \right). \]

We have completed the model derivation.
References


