

## Article

# Assessing Investor Belief: An Analysis of Trading for Sustainable Growth of Stock Markets

Yan Han <sup>1</sup>, Xue-Feng Shao <sup>2</sup>, Xin Cui <sup>3</sup>, Xiao-Guang Yue <sup>4,\*</sup>, Kelvin Joseph Bwalya <sup>5,\*</sup> and Otilia Manta <sup>6</sup>

<sup>1</sup> School of Humanities and Social Science, Beijing Institute of Technology, Beijing 100081, China; hanyan@bit.edu.cn

<sup>2</sup> Business School, University La Trobe Sydney Campus, Sydney, NSW 2000, Australia; x.shao@latrobe.edu.au

<sup>3</sup> Business School, University of International Business and Economics, Beijing 100029, China; cui.x@uibe.edu.cn

<sup>4</sup> Rattanakosin International College of Creative Entrepreneurship, Rajamangala University of Technology Rattanakosin, Nakon Patom 73170, Thailand

<sup>5</sup> School of Consumer Intelligence and Information System, University of Johannesburg, APK Campus, Johannesburg 2006, South Africa

<sup>6</sup> Romanian Academy, Center for Financial and Monetary Research-Victor Slăvescu, 010071 Bucharest, Romania; otilia.manta@univath.ro

\* Correspondence: xgyue@foxmail.com (X.-G.Y.); kbwalya@uj.ac.za (K.J.B.)

Received: 28 August 2019; Accepted: 8 October 2019; Published: 11 October 2019



**Abstract:** Investors' beliefs are the driving force behind the trading of stocks and, hence, sustainable stock returns. Although investors' beliefs are usually unobservable, this study develops a new approach to estimate investors' beliefs. Following well-established rational learning and market microstructure models, it is assumed that informed traders submit market orders according to their beliefs, whereas market makers/uninformed traders make Bayesian inferences about the informed traders' private signals after observing the total order flows. By fitting intraday transaction data to this model, we can estimate the daily belief uncertainties of informed and uninformed investors; this estimation is performed on S&P 500 stocks. The belief parameters estimated by this approach have incremental explanatory power to bid-ask spreads. The findings show that market makers' belief uncertainty plays a more important role in determining sustainable stock returns than informed traders'. Implications of these findings include: (a) the larger market maker group is influencing the market trends; (b) this dominant group is making decisions based on diverse types of data; and (c) increased understanding of the diversity of belief parameters may facilitate strategies to enhance sustainable returns, however, stock trading is still significantly influenced by emotive factors worthy of further research.

**Keywords:** investors' beliefs; information asymmetry; liquidity; stock trading; stock returns

## 1. Introduction

Bubbles are common in financial markets. When these bubbles are eventually burst, it often results in catastrophic consequences for the markets, and the negative impacts often last for several years. In this sense, sustainable growth of financial markets depends crucially on preventing the formation of bubbles. The numerous causes of bubbles can include the unrealistic, untrue, and unsustainable beliefs of traders, so assessing these beliefs is a logical starting point for investigating sources of bubbles. However, traders' beliefs are not directly observable. In this paper, we develop a method to indirectly measure traders' beliefs based on observed series of prices and orders.

One reason why people trade stocks is that they believe current stock prices differ from what they should be, or their true value. The more certainty traders have in their beliefs, the more aggressively they will trade, although they may be concerned that they know less about the stocks than other traders. And this uncertainty is very likely to affect trading. However, assessing traders' beliefs is important for us to understand their trading behavior. Unfortunately, traders' beliefs are seldom directly observable, and finance researchers have to rely on proxies—accounting report quality, analyst coverage, operation opacity, etc.—to measure traders' beliefs [1]. It would be logical to expect that the economic activities behind these proxies are correlated with some aspects of traders' beliefs, but it is hard to enumerate all factors affecting traders' beliefs. In this paper, we use a different approach to measure traders' beliefs.

While traders trade on their beliefs, an individual's trading activity may reveal his/her belief to other traders [2]. Accordingly, traders update their beliefs as they participate in trading. The updates may lead to new or modified trades, and the new trading may lead to newer beliefs, and so forth. In this way, the trading process reflects the evolution of traders' beliefs, which makes it possible to infer beliefs from trading. There are informed traders, market makers, and noise traders [3,4]. Traders' beliefs follow normal distributions. Initially, informed traders have private signals over market makers. Then, informed traders place orders of optimal sizes according to their beliefs. These orders are mingled with those of noise traders. Market makers and informed traders observe the order flows, make inferences, update their beliefs, and then respectively adjust prices and make further trading.

Traders' belief parameters can shed new insights into some empirical issues. Information asymmetry may have contrary effects on traders' actions [5]. Information asymmetry hinders traders in trading stock because of the adverse selection problem it creates, while it also makes price more informative due to the informed trading it induces. With the help of our estimates of traders' belief, we demonstrate that the second effect dominates. Whilst information asymmetry plays some role in determining stock prices, market makers' belief uncertainty seems to be a more important factor in terms of the rise and fall of stock prices. Therefore, our results show that to maintain a sustainable growth of financial markets, it is important to increase the information environment of the relatively less informed investors.

There is a large body of literature on how traders update their beliefs during trading [3,6–9]. Early rational expectation studies established that traders determine their optimal positions on assets according to their own beliefs and a public information indicator—stock prices [6]. In other words, current price conveys other traders' previous beliefs, and new traders consider this information when they trade. Bayesian learning research further explains how traders' beliefs change upon observing new information [10]. In these models, traders/agents follow Bayesian rules to update their beliefs when new information is received, and the rules specify the weight traders put on various sources of information. However, both rational expectation and Bayesian learning literature often emphasize finding and characterizing equilibria. Unsurprisingly, these models are not specific on the high frequency details of traders' belief-action interactions. The lack of these details creates difficulties in making reliable econometric inferences.

The models in market microstructure literature often provide detailed analysis on how price changes as a response to incoming orders [11]. To the extent that prices reflect market makers' beliefs, these models deal with traders' belief-action interactions at high-frequency. In addition, these models often divide traders into different groups according to their relative informational advantages [12]. Accordingly, market microstructure models can depict much finer pictures of trading and traders' beliefs. However, in Bayesian learning terms, the assumptions of traders' beliefs in market microstructure models are simplistic. For example, new information is often assumed to cause the same amount of change in traders' beliefs over time, whereas a Bayesian trader's belief change should vary with the frequency and accumulation of new information [10].

Our estimates of traders' belief parameters can well explain the stocks' bid-ask spreads. Moreover, we show that market makers' belief uncertainty is a key factor determining daily stock returns. Belief

uncertainty depresses traders' willingness to hold the stock, leading to lower contemporaneous stock return and higher future return.

This study contributes to understanding of the topic by providing a new set of traders' belief parameters. Conventional belief proxies are on a quarterly or even annual basis. However, as highlighted by Tannous et al. [13], traders' beliefs have large variations even over very short periods. So conventional belief proxies cannot explain fluctuations within quarters. Furthermore, conventional belief proxies do not distinguish between beliefs of different traders. Despite some common factors affecting all traders' beliefs, different traders' beliefs are likely to vary because of their differing information sources, analytical and interpreting skills.

The approach of this study is not based on what correlates with traders' beliefs, but to infer belief parameters from their trading; specific objectives are: (a) to incorporate both Bayesian learning and market microstructure models to develop a new model of how traders' beliefs are updated during trading; (b) to estimate the diversity of traders' belief parameters; and (c) investigate relationships of trader uncertainty, informational equality, and incorporation rates.

## 2. Methods

### 2.1. The Traders' Belief Settings

Following the literature, this model divides all traders into three groups: market makers, informed traders, and noise traders. Because market makers are just the mechanism to set price, we assume market makers have the same belief, whereas, informed traders and noise traders have different beliefs. At the beginning of trading, market makers acquire a common signal  $c$  about the firm value per share  $v$ :  $c = v + \epsilon_c$ ,  $\epsilon_c \sim N(0, \tau_c^{-1})$ . Therefore, the identical belief shared by all market makers is

$$v|c \sim N(c, \tau_c^{-1}). \quad (1)$$

Equation (1) writes the belief's variance in the form of its reciprocal i.e.,  $\tau_c^{-1}$ . This accords with the literature, which uses the variance's reciprocal  $\tau_c$  as the signal/belief value. Unlike market makers who share an identical belief, informed and noise traders have differing beliefs. We use  $i \in [0, I]$  and  $h \in [0, 1]$  to designate the informed traders and noise traders, respectively. All informed traders also know the common signal  $c$ . Besides  $c$ , each informed trader  $i$  also acquires a private signal  $s_i = v + \epsilon_{s,i}$ ,  $\epsilon_{s,i} \sim N(0, \tau_s^{-1})$ . The private signals have the same precision, i.e.,  $\tau_s^{-1}$ , among informed traders. The diversity of private signals comes from  $s_i$ , which is only known to informed traders  $i$ . According to the Bayesian theorem, after observing signals  $c$  and  $s_i$ , an informed trader  $i$  believes that

$$v|c, s_i \sim N\left(\frac{\tau_c c + \tau_s s_i}{\tau_c + \tau_s}, (\tau_c + \tau_s)^{-1}\right). \quad (2)$$

Equation (2) demonstrates the useful property of Bayesian learning with normally distributed beliefs. That is, both pre- and posterior prior beliefs follow normal distributions. The posterior mean is the "precision-weighted" average of the prior belief and the new signal; and the posterior precision is the sum of the prior's and new signal's precisions. Vives [10] gives a formal proof of these properties.

Having initiated the beliefs of market makers and informed traders, we start investigating the trading process. At the beginning of each round  $t$  of trading, market makers set the stock price  $p_t$ , which is equal to their mean belief. In reality, the price of a particular stock at a particular time is not quoted as a single value, but as a set of bid prices and a set of ask prices. However, market microstructure literature has shown that the midpoint of the best bid and ask prices reflects the current beliefs of market makers [14,15], in this instance, the stock price  $p_t$  in our model should be viewed as the midpoint of bid-ask price at time  $t$ .

After market makers set price  $p_t$  at the beginning of trading round  $t$ , informed and noise traders submit orders anonymously. To equate the buy and sell orders, market makers set a new price  $p_{t+1}$ ,

which starts trading round  $t + 1$ . In other words, we do not assume the equilibrium price to be established instantly, but rather to be found in a trial and error process. The way equilibrium price is established distinguishes two schools of thought. The earlier literature adopts auction models [14–17]. In auction models, there are no rounds of trading. All traders submit orders at the same time. Then an equilibrium price is established by maximizing the transaction volume when buyers and sellers are matched. Although theoretically insightful, the auction models are distant from what happens in the real stock markets, where trades occur second by second, and no price can be nominated as being in equilibrium. To shed more light on the price dynamics in real stock markets, more recent literature begins to examine how prices, trades, and beliefs interact with each other in a sequential setting [14–17]. This recent literature is not interested in equilibrium prices, which are not empirically well defined or readily testable. On the contrary, this school of thought focuses on the trade-price process through which an equilibrium can be achieved eventually. Our model is consistent with this recent approach.

## 2.2. The Update of Beliefs

Since our model allows for rounds of trading, traders update their beliefs after observing each round of trading. New beliefs stimulate new trading, and so on and so forth. We now use a mathematical model to demonstrate this belief updating process.

In round  $t$ , let  $x_{i,t}$  and  $y_{h,t}$  denote the order size of informed trader  $i$  and noise trader  $h$ , respectively. Therefore, the total order size is  $q_t = \int_0^I x_{i,t} di + \int_0^1 y_{h,t} dh$ . Notice that  $q_t$  is known to all market participants once the orders are submitted and executed. Therefore,  $q_t$  serves as a new signal to all traders.

Noise traders do not trade on the basis of any information, hence the name “noise”. It follows that the order size of each noise trader should be independent. Let us further assume that noise traders independently choose their order size from  $N(0, \sigma_n^2)$ , i.e.,  $\int_0^1 y_{h,t} dh \sim N(0, \sigma_n^2)$ .

Informed traders trade on their own beliefs. To determine how their beliefs affect their trades, we assume that informed traders are utility maximizers. Let informed traders’ utility function be  $U(w) = -\exp(-\rho w)$ , where  $\rho$  is the trader’s risk aversion coefficient and  $w$  is the wealth. For informed trader  $i$  who believes  $v \sim N(\bar{v}_t, \tau_t^{-1})$ , a standard mathematical practice in asset pricing literature can show that

$$x_{i,t} = \frac{\tau_t}{\rho} (\bar{v}_t - p_t). \quad (3)$$

According to Equations (2) and (3), informed trader  $i$ ’s order size in round  $t = 1$  is

$$x_{i,1} = \frac{\tau_c + \tau_s}{\rho} \left( \frac{\tau_c c + \tau_s s_i}{\tau_c + \tau_s} - p_1 \right) = \frac{1}{\rho} [\tau_c c + \tau_s s_i - (\tau_c + \tau_s) p_1]. \quad (4)$$

When market makers and informed traders observe the total order  $q_1$ , which includes  $x_{i,1}$  and orders from noise traders, they infer  $v$  based on  $q_1$ . For market makers, although they do not know  $s_i$ , they do know that  $s_i | v \sim N(v, \tau_s^{-1})$  and consequently the following two equations:

$$x_{i,1} | v, c \sim N \left( \frac{1}{\rho} [\tau_c c + \tau_s v - (\tau_c + \tau_s) p_1], \frac{\tau_s}{\rho^2} \right). \quad (5)$$

$$q_1 | v, c \sim N \left( \frac{1}{\rho} [\tau_c c + \tau_s v - (\tau_c + \tau_s) p_1], \frac{I \tau_s}{\rho^2} + \sigma_n^2 \right). \quad (6)$$

Informed traders also want to infer  $v$  based on  $q_1$  because they only have their own imprecise private signals, and knowledge of other informed traders’ private signals assists in enhancing their

belief precision. To show how informed traders make inferences, we first write informed trader  $i$ 's conditional belief of  $q_1$  as

$$q_1|v, c, s_i \sim N\left(\frac{I}{\rho}[\tau_c c + \tau_s v - (\tau_c + \tau_s)p_1], \frac{I\tau_s}{\rho^2} + \sigma_n^2\right). \quad (7)$$

Then we can apply the Bayesian theorem to Equations (1) and (6) and calculate the market makers' posterior belief  $v|c, q_1$  at the end of round 1 as:

$$\tau(v|c, q_1) = \tau_c + \frac{\left(\frac{I}{\rho}\tau_s\right)^2}{\frac{I\tau_s}{\rho^2} + \sigma_n^2} = \tau_c + \tau_q. \quad (8)$$

$$\mathbb{E}(v|c, q_1) = \frac{\tau_c}{\tau_c + \tau_q}c + \frac{\tau_q}{\tau_c + \tau_q}\left(\frac{\rho}{I\tau_s}q_1 - \frac{\tau_c}{\tau_s}c + \frac{\tau_c}{\tau_s}p_1 + p_1\right). \quad (9)$$

Given that market makers always set prices based on their own beliefs, stock price in the beginning of trading round  $t = 2$  is  $p_2 = \mathbb{E}(v|c, q_1)$ .

Similarly, applying the Bayesian theorem to Equations (2) and (7), we can determine the informed trader  $i$ 's posterior belief  $v|c, s_i, q_1$  as:

$$\tau(v|c, q_1) = \tau_c + \tau_s + \frac{\left(\frac{I}{\rho}\tau_s\right)^2}{\frac{I\tau_s}{\rho^2} + \sigma_n^2} = \tau_c + \tau_s + \tau_q. \quad (10)$$

$$\mathbb{E}(v|c, q_1) = \frac{\tau_c\tau_c c + \tau_s s_i}{\tau_c + \tau_s + \tau_q} + \frac{\tau_q}{\tau_c + \tau_s + \tau_q}\left(\frac{\rho}{I\tau_s}q_1 - \frac{\tau_c}{\tau_s}c + \frac{\tau_c}{\tau_s}p_1 + p_1\right). \quad (11)$$

The posterior beliefs of market makers and informed traders specified by Equations (8) through (11), in turn determine  $p_2$  and  $q_2$  in round  $t = 2$ . In round  $t = 2$ , similar belief updating occurs, leading to the next price  $p_3$  at the start of the next round of trading  $q_3$ . By induction, at the end of round  $t$ , we have

$$v|c, q^{(t)} \sim N\left(\frac{\tau_c c + \tau_q H_t}{\tau_c + t\tau_q}, \frac{1}{\tau_c + t\tau_q}\right). \quad (12)$$

$$v|c, s_i, q^{(t)} \sim N\left(\frac{\tau_c c + \tau_s s_i + \tau_q H_t}{\tau_c + \tau_s + t\tau_q}, \frac{1}{\tau_c + \tau_s + t\tau_q}\right). \quad (13)$$

where  $q^{(t)} = \{q_1, q_2, \dots, q_t\}$  and

$$H_t = \sum_{k=1}^t \left(1 - \frac{\tau_q}{\tau_s}\right)^{t-k} \left[\frac{\rho}{I\tau_s}q_k + \left(1 + \frac{(k-1)\tau_q}{\tau_s}\right)p_k + \frac{\tau_c}{\tau_s}(p_k - c)\right]. \quad (14)$$

$$\tau_q = \frac{(I\tau_s)^2}{I\tau_s + \rho^2\sigma_n^2}. \quad (15)$$

Now considering Equations (12) and (13) in more detail; as these define the updated beliefs of market makers and informed traders, respectively. Notice that we assume all traders' beliefs to follow normal distributions, and all normal distributions can be determined by two parameters, i.e., mean and variance. This means that we can interpret the belief updating process by looking at how the mean and precision (reciprocal of variance) of beliefs evolve over time.

At the end of round  $t$ , market makers' and informed traders' belief precisions are  $\tau_c + t\tau_q$  and  $\tau_c + \tau_s + t\tau_q$ , respectively. After each round, the belief precisions of both market makers and informed traders are enhanced by the same amount  $\tau_q$ .

At the end of round  $t$ , market makers' and informed traders' belief means are the precision-weighted average of all information that they have. Specifically, Equations (12) and (14) demonstrate that market makers have two pieces of information: the initial common signal  $c$  and the history of  $t$  rounds of orders. The precision of  $c$  is a constant  $\tau_c$ . The precision of each round of order is also a constant, i.e.,  $\tau_q$ . Therefore, the relative precisions, hence the weights, of the common signal  $c$  and of each order are  $\tau_c/(\tau_c + t\tau_q)$  and  $\tau_q/(\tau_c + t\tau_q)$ , respectively.

Equation (14) describes the history of trading and translates it into a new signal. Although the right-hand side of Equation (14) appears complex, it can be viewed as two parts. The first part consists of the terms in the round bracket. It is easy to prove that  $0 < 1 - \tau_q/\tau_s < 1$ . Thus, the first part serves as a *discount factor*, which discounts each round of orders to its present value. This discount term gives more weight to more recent orders. In other words, the more recent an order is, the more informative it is. This property makes intuitive sense because the informativeness of orders depends on informed traders' belief precisions. Over time, informed traders' beliefs become more precise, leading to a more informative order flow.

The second part, in the square bracket of Equation (14) can be viewed as a *translation term*, which translates the total order size  $q_k$  in round  $k$  into new signals. The translation term of Equation (14) is similar, with the equation defining informed trader  $i$ 's order size  $x_{i,t}$  in Equation (3). To show this point, let us rearrange Equation (3) in the following form:

$$\bar{v}_t = \frac{\rho}{\tau_t} x_{i,t} + p_t. \quad (16)$$

In the above form of Equation (16), it demonstrates that Equation (3) is actually also a *translation function*, which translates informed trader's order size  $x_{i,t}$  to a mean belief. Equation (14) is much more complicated than Equation (16) simply because the former adjusts for the information precision of all rounds of orders.

### 2.3. Relationship Between Prices and Order Flows

As we have shown that beliefs, prices, and orders interact with each other, we will now proceed to establish the mathematical relationships between them. First, the Equations (12) and (13) can be rewritten as:

$$p_{t+1}|c, q^{(t)}, p^{(t)} = \frac{\tau_c c + \tau_q H_t}{\tau_c + t\tau_q} + e_{p,t+1}, \quad (17)$$

where

$$e_{p,t+1} \sim N\left(0, \frac{1}{\tau_c + t\tau_q}\right), \quad (18)$$

and

$$q_{t+1}|v, c, q^{(t)}, p^{(t+1)} = \frac{I}{\rho} (\tau_c c + \tau_s v + \tau_q H_t - (\tau_c + \tau_s + t\tau_q)p_{t+1}) + e_{q,t+1}, \quad (19)$$

where

$$e_{q,t+1} \sim N\left(0, \frac{I\tau_s}{\rho^2} + \sigma_n^2\right). \quad (20)$$

Rearranging Equations (17) and (19), we have

$$q_t = \left(1 + \frac{\tau_q}{\tau_s}\right) q_{t-1} - \frac{I}{\rho} [\tau_c + \tau_s + (t-1)\tau_q] (p_t - p_{t-1}) + \epsilon_t, \quad (21)$$

where

$$\epsilon_t \sim N\left(0, \frac{I^2}{\rho^2} \left( \frac{2\tau_s^2}{\tau_q} + \frac{[\tau_c + (t-2)\tau_q]\tau_q^2}{\tau_s^2} \right)\right). \quad (22)$$

Equation (21) includes the price series  $\{p_t\}$ , order series  $\{q_t\}$ , and a zero-mean normally distributed error term  $\{\epsilon_t\}$ , along with some scalar parameters. In this form, Equation (21) demonstrates the mathematical relation among beliefs, prices, and orders. Moreover, Equation (21) can also be viewed as a regression function of the next order on the previous orders and prices. This regression function has two main implications, which are discussed below.

First, Equation (21) illustrates a positive autocorrelation among order series  $\{q_t\}$ , which is consistent with empirical results in prior studies (see Van Ness et al. [18] for a brief summary of these results). Second, the coefficient of  $q_{t-1}$  is  $1 + \tau_q/\tau_s$ , which is strictly larger than 1. However, it does not imply the order series  $\{q_t\}$  will explode over time. That is because the second term of the right-hand side of Equation (21) exerts an opposite effect on the next order. For example, let us assume the previous order is a buy order, i.e.,  $q_{t-1} > 0$ . Apparently, this buy order moves the price up, i.e.,  $p_t - p_{t-1} > 0$ . Therefore, the second term is negative, causing the next order  $q_t$  to be smaller.

#### 2.4. Data and the Belief Estimation Procedure

Since Equation (21) can be viewed as a regression, and the variables  $p_t$  and  $q_t$  are observable, we can use this regression function to estimate its unobserved parameters:  $\tau_s$ ,  $\tau_c$ ,  $\tau_q$ , and  $\rho/I$ . The first three parameters are of great economic interest as they represent the precision of informed traders' private signals, the precision of common signals, and the precision of the signal conveyed by order flows, respectively.

We apply this estimation to the data of S&P 500 index constituent stocks over the period between 1 January 2003 and 31 December 2011. To avoid the impacts of S&P 500 index constituent changes, we have excluded the first month for newly added stocks and the last month for newly removed stocks.

The estimation requires detailed series of prices and orders. This paper uses the NYSE's Trade and Quote (TAQ) database, which provides order-by-order trades and quotes data. Before the estimation, we clean TAQ raw data in the following ways. First, trades that occur at the same time and price are aggregated as one trade. Second, earlier microstructure literature routinely adjusts the timestamps in pre-2000 TAQ data [19]. It is unclear whether similar timestamp adjustments should be applied to post-2000 TAQ data [20]. Specifically, Bessembinder [21] and Boehmer and Kelley [22] do not adjust timestamps, whereas Brennan et al. [23] and Tannous et al. [13] do. In this paper, we use both time-adjusted and time-unadjusted data; however, as results are quantitatively similar, we only report results of time-unadjusted data.

We also removed obviously erroneous data entries from the raw data according to widely accepted principles [24,25]. These omitted trade data entries are: (1) out of time sequence order executions; (2) trades with an error or correction indicator; (3) trades that are not settled in standard ways; (4) trades that are associated with exchange acquisitions or distributions; (5) trades that are not preceded by a valid same-day quote; (6) trades that are executed at a price changed 50% from that of the previous trade. The removed quote data entries are: (1) quotes with negative bid or ask prices; (2) quotes with negative bid-ask spreads; (3) quotes with bid-ask spreads larger than \$5; (4) quotes with the midpoint of bid and ask prices changed by 50% from the previous same-day midpoint; and (5) quotes associated with trading halts or designated order imbalances.

After cleaning the raw data, we apply an estimation for each stock on each trading day, using a maximum likelihood method. To avoid local optima, we run the estimations 10 times with different initial values and keep the estimates that have the largest likelihood.

#### 2.5. Definitions and Estimates of Belief Parameters

The estimation generates estimates for four parameters:  $\tau_s$ ,  $\tau_c$ ,  $\tau_q$ , and  $\rho/I$ , although the last parameter is not of interest in this paper and is not reported. The first three parameters represent information precisions of informed traders' private signals, common signals, and the signal conveyed by order flows. To better interpret these three parameters, we define three variables. The first variable is belief uncertainty. It is actually the price-adjusted precision. The rationale for the adjustment is as



follows. Precisions are mathematical variations of prices, and thus stocks with higher prices necessarily have smaller precision values. This fact prevents comparisons of belief precisions among stocks with different price levels. Specifically, we define belief uncertainty by the following Equation:

$$\text{Belief uncertainty} = \frac{6 \times \tau^{-0.5}}{\text{Average transaction price during the day}}. \quad (23)$$

In Equation (23), informed traders'  $\tau$  is  $\tau_c + \tau_s$  and market makers' is  $\tau_c$ . Equation (23) uses average transaction price during the day to proxy for true value of the stock. As that precision  $\tau$  is the reciprocal of belief variance and we assume the beliefs are normally distributed, then Equation (23) uses  $6 \times \tau^{-0.5}$  to represent the range of likely prices. Belief uncertainty is a measure of how uncertain traders are about the signals. The larger the belief uncertainty is, the more uncertain traders feel.

The second variable we define is information equality, which measures the informational advantage of informed traders over market makers. This variable is defined as:

$$\text{Information equality} = \frac{\text{Informed traders' belief uncertainty}}{\text{Market makers' belief uncertainty}}. \quad (24)$$

Information equality can be viewed as an inverse measure of information asymmetry, for the following reasons. When the information asymmetry is severe, the informed traders are much more certain than market makers. In this case, the numerator of information equality is relatively smaller, leading to a smaller information equality.

The third variable we define is information incorporation rate. This variable focuses on how aggressively market makers adjust prices after observing order flows. The level of aggression of market adjustment is dependent on how much weight market makers put on the signals conveyed by order flows. If the signal conveyed by order flows is more informative relative to their own common signal  $c$ , i.e.,  $\tau_q$  is relatively larger than  $\tau_c$ , then market makers put more weight on the order flow signal and adjust prices more aggressively. Accordingly, we define information incorporation rate as follows:

$$\text{Information incorporation rate} = \frac{\tau_q}{\tau_c}. \quad (25)$$

In this instance, the aggressiveness of market makers' adjustments of prices should decrease over time, because after observing more and more rounds of orders they become less and less uncertain, thus putting less and less weight on the signal of the newest orders. In this context, one should view the variable information incorporation rate as a proxy for the decreasing aggressiveness of price adjustments. Another property of this variable is that it also measures how quickly prices can fully incorporate all information in the market.

### 3. Results and Discussion

#### 3.1. Belief Parameter Estimates over Time

Prior to empirical testing of belief estimates we summarized the annual estimates over the study period (see Table 1).

Table 1 summarizes the estimates of belief parameters during each year. On average, informed traders' belief uncertainty is 0.0198, which means that if the unknown true stock price is \$100, then informed traders generally believe the stock price should be between \$99.01 and \$100.99. Similarly, market makers' average belief uncertainty is 0.0774, indicating that they generally believe that the stock price should be between \$96.13 and \$103.87. Both informed traders' and market makers' belief uncertainty is larger in the years 2008 and 2009, probably due to the financial crisis. The average informational equality is 0.4329, meaning that market makers feel half as certain as informed traders. The informational equality shows large variations with some stocks having informational equality as low as 0.09. In other words, informed traders' beliefs are 10 times more precise than market makers.



However, some stocks have informational equality close to 1, indicating little information asymmetry; this is likely to be true for large and closely monitored firms in the USA. The information incorporation rate also shows great variation. For some stocks the incorporation rate is almost 0, so trading in these stocks hardly carries any new information.

**Table 1.** Summary of four belief parameter estimates, for a nine-year period.

	Variable	Year	Mean	Std	Q1	Median	Q3
(1)	Informed trader's belief uncertainty	2003	0.0188	0.0433	0.0060	0.0097	0.0197
(2)		2004	0.0165	0.0354	0.0053	0.0086	0.0188
(3)		2005	0.0170	0.0453	0.0050	0.0081	0.0168
(4)		2006	0.0160	0.0460	0.0046	0.0076	0.0158
(5)		2007	0.0153	0.0281	0.0036	0.0071	0.0175
(6)		2008	0.0256	0.0514	0.0045	0.0092	0.0247
(7)		2009	0.0295	0.0541	0.0051	0.0118	0.0360
(8)		2010	0.0203	0.0367	0.0035	0.0092	0.0274
(9)		2011	0.0188	0.0337	0.0031	0.0080	0.0249
(10)		All	0.0198	0.0426	0.0046	0.0087	0.0217
(11)	Market maker's belief uncertainty	2003	0.1185	0.1972	0.0363	0.0630	0.1157
(12)		2004	0.0908	0.1606	0.0298	0.0500	0.0873
(13)		2005	0.0899	0.1789	0.0283	0.0462	0.0800
(14)		2006	0.0734	0.1483	0.0255	0.0419	0.0684
(15)		2007	0.0492	0.0898	0.0152	0.0350	0.0553
(16)		2008	0.0729	0.1150	0.0197	0.0479	0.0853
(17)		2009	0.0877	0.1675	0.0192	0.0521	0.0905
(18)		2010	0.0610	0.1227	0.0110	0.0379	0.0631
(19)		2011	0.0536	0.1039	0.0097	0.0351	0.0589
(20)		All	0.0774	0.1481	0.0235	0.0442	0.0769
(21)	Informational equality	2003	0.3192	0.3565	0.0966	0.1052	0.4245
(22)		2004	0.3629	0.3753	0.0966	0.1714	0.7152
(23)		2005	0.3623	0.3686	0.0988	0.1717	0.6639
(24)		2006	0.3907	0.3743	0.0991	0.1731	0.8735
(25)		2007	0.4744	0.3834	0.1518	0.2140	0.9984
(26)		2008	0.4624	0.3763	0.1447	0.2087	0.9856
(27)		2009	0.5037	0.3852	0.1714	0.3215	0.9992
(28)		2010	0.5115	0.3840	0.1714	0.3571	0.9993
(29)		2011	0.5058	0.3810	0.1714	0.3462	0.9990
(30)		All	0.4329	0.3826	0.1025	0.2038	0.9903
(31)	Information incorporation rate	2003	0.6183	17.7589	0.0002	0.0020	0.0105
(32)		2004	0.6331	19.2780	0.0001	0.0020	0.0115
(33)		2005	0.6211	12.8847	0.0002	0.0021	0.0153
(34)		2006	0.6779	15.6389	0.0001	0.0020	0.0193
(35)		2007	0.4294	15.7945	0.0000	0.0014	0.0387
(36)		2008	0.2187	3.8260	0.0001	0.0022	0.0872
(37)		2009	0.3221	5.5363	0.0000	0.0018	0.0497
(38)		2010	0.6071	21.7317	0.0000	0.0025	0.0558
(39)		2011	0.5660	14.8554	0.0000	0.0034	0.0841
(40)		All	0.5209	15.2114	0.0000	0.0020	0.0315

The statistics listed are mean and standard deviation values, of the 1st quintile, median, and 3rd quintile.

### 3.2. Characterizing Belief Parameters

The belief parameters estimated by our model, separately depict the belief uncertainties for informed traders and market makers. These two groups of traders' beliefs are all affected by the firm's information environment which, in turn, is determined by factors such as the quality of the firm's financial reporting, operation opacity and analyst's coverage, etc.; this implies that belief parameters may exhibit correlations.

Table 2 summarizes these correlation coefficients. The correlation coefficient between informed traders' and market makers' belief uncertainty is 0.43, indicating that there may be common underlying factors that affect the information environment for all traders. Information equality should be positively

correlated with informed trader's belief uncertainty, and negatively correlated with market makers' belief uncertainty, which the table values confirm.

**Table 2.** Correlations among the belief parameters.

Variable	Ifmd. Belief Uncertainty	MM. Belief Uncertainty	Info. Equality	Incp. Rate
Ifmd. belief uncertainty	1.00			
MM. belief uncertainty	0.43 ***	1.00		
Info. equality	0.28 ***	−0.24 ***	1.00	
Incp. rate	−0.01 ***	0.19 ***	−0.03 ***	1.00

Ifmd. belief uncertainty is informed traders' belief uncertainty, MM. belief uncertainty is market makers' belief uncertainty, Info. equality is the informational equality, and Incp. rate is the information incorporation rate. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

Information incorporation rate is positively correlated with market makers' belief uncertainty, and negatively correlated with informed traders' belief uncertainty. This relationship makes sense, because with larger belief uncertainty, market makers put more weight on new information, adjusting price more aggressively. Then when informed traders' belief uncertainty is larger, they trade less aggressively, making order flows less informative, which leads to a lower information incorporation rate. However, the correlation between informed traders' belief uncertainty and information incorporation rate is low, though still significant. This suggests that informed traders are playing a less important role than market makers in the market. Similarly, information incorporation rate is also negatively correlated with informational equality, because when market makers feel equal with informed traders in terms of information advantage, market makers are less keen to seek new information, thus adjusting prices less aggressively. However, comparing the correlation with information incorporation rate, market makers' belief uncertainty has a much stronger correlation than informed traders' belief uncertainty or informational equality does. These results indicate that the level of uncertainty of market makers is more important in determining the aggressiveness of price adjustment.

Besides the correlations among belief parameters, a firm's information environment may also induce persistence in belief parameters, because the information environment is likely to be stable over a long-time horizon. Thus, daily belief parameters should be persistent with time. To investigate this persistence, we subjected the belief parameters to regression in relation to their lags. Table 3 summarizes the regression results.

**Table 3.** Persistence of the belief parameters.

	Ifmd. Belief Uncertainty	MM. Belief Uncertainty	Info. Equality	Incp. Rate
	(1)	(2)	(3)	(4)
Intercept	0.0072 *** (0.0002)	0.0144 *** (0.0003)	0.3177 *** (0.0034)	0.4715 *** (0.0211)
Lag 1	0.1429 *** (0.0043)	0.1858 *** (0.0035)	0.0588 *** (0.0013)	0.0602 *** (0.0177)
Lag 2	0.1351 *** (0.0043)	0.1659 *** (0.0031)	0.0557 *** (0.0013)	0.0280 *** (0.0090)
Lag 3	0.1312 *** (0.0040)	0.1570 *** (0.0034)	0.0485 *** (0.0013)	0.0305 *** (0.0109)
Lag 4	0.1251 *** (0.0042)	0.1572 *** (0.0033)	0.0480 *** (0.0012)	0.0257* (0.0138)
Lag 5	0.1296 *** (0.0040)	0.1562 *** (0.0033)	0.0454 *** (0.0012)	0.0414 * (0.0236)
R <sup>2</sup>	22.01%	43.00%	1.94%	3.76%
Obs	890,771	890,771	890,771	890,771

This table summarizes results from Fama and MacBeth [26] regressions of belief parameters on their lag values. The dependent variables in Columns 1 through 4 are informed traders' belief uncertainty, market makers' belief uncertainty, informational equality, and information incorporation rate, respectively. The regressions are run on daily data. Standard errors are in parentheses. \*, \*\*, and \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively.

When dealing with panel data, Petersen [27] recommends using the Fama and MacBeth [26] regression for asset pricing applications, and using standard errors clustered within firm and/or time for corporate finance applications.

Table 3 shows that both traders' belief uncertainty and informational equality are highly persistent, at least up to lag 5. This persistence reflects the stable nature of a firm's information environment. Information incorporation rate is also persistent, but to a lesser extent. The coefficients of lag 4 and 5 are only marginally significant.

Although a firm's stable information environment can cause persistence in belief parameters, informational events happen frequently to firms. With the help of rapidly expanding information technologies and the internet, an event can be known to traders across the globe within a second, changing millions of traders' beliefs. So, it is logical to expect large variations in daily belief parameters. To measure the variation of belief parameters, we first calculate the standard deviation and mean of each belief parameter for each stock within each month. Then, we divide the standard deviation by the mean, so the result measures the variation of the belief parameter relative to its average value within a month. We call this result monthly variation of the belief parameter.

Table 4 reports the distribution of monthly variation for each belief parameter. The median monthly variations of informed traders' and market makers belief uncertainty are 0.97 and 0.59, respectively. These variations are large. Within a month, an average trader's belief uncertainty is very likely to double from day to day, as the standard deviation of the belief uncertainty is close to its mean. Notably, informed traders' belief uncertainty has larger variation than that of market makers' belief. Informational equality and information incorporation rate also have large variation. Additional background analyses (not summarized here) showed the variations within quarters are even larger. The large variations demonstrate that the information environment proxies used in prior studies—for example, asset tangibility, accounting accruals, cash flow volatility, and probability of informed trading (PIN),—do not incorporate significant amounts of information. The belief parameters estimated by our model can provide finer measurements of the information environment.

**Table 4.** Monthly variation of belief parameters.

	5th	25th	50th	75th	95th
Informed traders' belief uncertainty	0.40	0.78	0.97	1.22	1.82
Market makers' belief uncertainty	0.33	0.47	0.59	0.80	1.84
Informational equality	0.56	0.76	0.89	1.04	1.31
Information incorporation rate	1.42	1.95	2.48	3.18	4.33

This table lists the 5th, 25th, 50th, 75th, and 95th percentiles of monthly variation for the belief parameters. For each stock in each month, the monthly parameter variation is defined as the standard deviation of daily parameter estimates within the month, divided by the mean of the daily parameter estimates within the month. Each stock has one monthly variation for each month and the table summarizes the distribution of monthly parameter variations across all stocks in all months.

### 3.3. Traders' Belief Uncertainty and Bid-Ask Spread

Prior studies have established that bid-ask spreads are the compensations for market makers providing liquidity. The adverse selection problem facing market makers is one of the determinants of bid-ask spreads [14–17]. Traders' belief uncertainty and informational equality can directly measure this adverse selection problem. Specifically, when market makers' belief uncertainty is larger, they are more concerned about adverse selection. Thus bid-ask spreads should be wider. When informational equality is lower, market makers believe that informed traders know the firm much more precisely than they do. Then, adverse selection is more acute for market makers and bid-ask spreads should be wider. As for informed traders' belief uncertainty, besides its effect captured in informational equality, it has another effect on bid-ask spreads. Informed trading makes price informative, causing market makers' belief to become more precise. Informed traders' belief uncertainty should affect bid-ask spreads in the same direction as market makers'. Therefore, the first hypothesis that this paper tests is:

**Hypothesis 1.** *Both informed traders' and market makers' belief uncertainty are positively related with bid-ask spreads, whereas informational equality is negatively related with bid-ask spreads.*

To test H1, we have undertaken regression of percentage bid-ask spreads on informed traders' and market makers' belief uncertainty and informational equality. Prior studies have shown that some stock attributes affect bid-ask spreads [1,13,28,29]. For example, market makers set narrower spreads for stocks with higher prices, because the fixed processing costs can be spread over a higher value. When the volatility of price is larger, market makers demand a larger compensation for providing liquidity, leading to a wider spread. Larger firms are more transparent, so the adverse selection problem is less severe, and market makers set narrower spreads. For stocks with more active trading, market makers find it easier to rebalance their inventory, so the spreads can be narrower. These variables are the control variables in the regression.

In Column 1 of Table 5, market makers' belief uncertainty is positively and significantly related to bid-ask spreads, which is consistent with the hypothesis that when market makers feel more uncertain, they set wider spreads. In Column 2, when adding informational equality as an additional explanatory variable, the coefficient of market makers' belief uncertainty remains positive and significant, while informational equality is negatively and significantly related to bid-ask spreads; this latter finding indicates that information asymmetry leads to wider spreads. In Column 3, we further add informed traders' belief uncertainty as an explanatory variable. The signs and significance of coefficients of market makers' belief uncertainty and informational equality remain unchanged. Informed traders' belief uncertainty is positively and significantly related to bid-ask spreads, supporting the hypothesis that informed trading leads to narrower spreads after controlling for information asymmetry.

**Table 5.** Traders' beliefs and bid-ask spreads.

	(1)	(2)	(3)
Intercept	0.2525 *** (0.0069)	0.2533 *** (0.0069)	0.2526 *** (0.0068)
MM. belief uncertainty	0.0117 *** (0.0007)	0.0101 *** (0.0007)	0.0084 *** (0.0007)
Info. equality		−0.0015 *** (0.0001)	−0.0023 *** (0.0002)
Ifmd. belief uncertainty			0.0170 *** (0.0026)
1/Price	0.8290 *** (0.0079)	0.8319 *** (0.0079)	0.8278 *** (0.0078)
Vol. tran. prc.	0.0459 *** (0.0008)	0.0459 *** (0.0008)	0.0459 *** (0.0008)
log(size)	−0.0099 *** (0.0003)	−0.0099 *** (0.0003)	−0.0099 *** (0.0003)
Turnover	−0.1915 *** (0.0104)	−0.1922 *** (0.0104)	−0.1915 *** (0.0105)
$R^2$	78.09%	78.12%	78.19%
Obs	893,782	893,782	893,782

This table reports results from Fama and MacBeth [26] regressions of percentage bid-ask spreads on traders' beliefs and stock characteristics. MM. belief uncertainty is market makers' belief uncertainty. Info. equality is informational equality. Ifmd. belief uncertainty is informed traders' belief uncertainty. 1/Price is the reciprocal of share price. Vol. tran. prc. is the volatility of transaction price. log(size) is the logarithm of market cap. Turnover is the stock's trading turnover. See Appendix A for detailed definitions of these variables. The regressions are run on daily data. Standard errors are in parentheses. \*, \*\*, and \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively.

Although informational equality, market makers' and informed traders' belief uncertainty are significantly associated with bid-ask spreads, the magnitude of their effects are different. In Column 3, one standard deviation change in market makers' belief uncertainty causes the spreads to change by 0.13 basis points, whereas the changes caused by one standard deviation change in informational

equality and informed traders' belief uncertainty are 0.09 and 0.07 basis points, respectively. Therefore, market makers belief uncertainty plays a much more important role in determining bid-ask spreads than does informed traders' belief uncertainty or informational equality.

The control variables show coefficient signs consistent with those in prior studies.  $1/\text{Price}$  and volatility of transaction price are positively associated with bid-ask spreads. Firm size ( $\log(\text{size})$ ) and turnover are negatively associated with bid-ask spreads.

### 3.4. Information Risk and Stock Returns

Is information asymmetry a pricing factor? This question has drawn much attention since Easley et al. [30]; Easley and O'Hara [31] argued that information asymmetry generates an "information risk", which commands a risk premium. However, this argument faces many challenges. Of particular note is the model of Lambert et al. [32], which shows that information asymmetry, per se, is not a pricing factor. What matters is traders' information precision, which determines their willingness to hold stock. Furthermore, Wang [5] developed a model in which information asymmetry has two contrary effects on traders' willingness to hold stock. On one hand, information asymmetry deters uninformed traders' willingness to hold the stock, on the other hand, information asymmetry implies more informed trading and a more informative price, which increase willingness. Empirical tests on the relationship between information risk and stock return are far from definitive, partly due to the lack of good proxies for information asymmetry. Belief parameters can provide more insight into this issue.

Equation (3) of our model echoes the key conclusion of Lambert et al. [32]. That is, traders tend to hold riskier assets if they feel more certain about the asset value. It follows that with lower information uncertainty, stock price should be higher, and contemporaneous stock return should also be higher. To test this hypothesis, we can analyze (by regression) daily stock returns on market makers' and informed traders' belief uncertainty on the same day. According to Wang [5], information asymmetry can both depress and enhance contemporaneous stock returns at the same time. To test which effect dominates, we can use informational equality as another explanatory variable. In summary, the second hypothesis is:

**Hypothesis 2.** *Market makers' belief uncertainty and informational equality are negatively related to the contemporaneous stock returns, whereas informed traders' belief uncertainty is positively related to the contemporaneous stock returns.*

Prior studies have also documented that order imbalances can explain contemporaneous and future stock returns. In part, because order imbalances drive market makers to adjust the price in the same direction [33–35]. According to our model, market makers' belief uncertainty should strengthen order imbalances' explanatory power, because an uncertain market maker will adjust price more aggressively towards the direction of order imbalances. We can test this prediction by including, in regressions, order imbalances and an interaction term between order imbalances and market makers' belief uncertainty.

The dependent variable in Column 1 of Table 6 is the open-close return. We construct this variable following Chordia and Subrahmanyam [34]. This open-close return can be distinguished from the conventional daily stock return in three ways. First, this stock return measures the price change during the trading sessions, whereas conventional daily stock return measures the change from the previous day's close price to the current day's close price. Thus, our open-close return excludes price/belief changes in the after-hours. Traders' belief uncertainty measures how uncertain traders feel during the trading sessions. Therefore, the open-close return is better matched with the explanatory variables than conventional daily stock returns. Second, the open and close prices are the first and last midpoint quoted prices associated with the first and last transactions of the day, excluding the opening batch auctions. Using the midpoint quoted prices can avoid the well-known bid-ask bounce bias when dealing with short horizon returns. Third, the open-close return is excess return,

in that we subtract market return from the open–close price change. Using excess return can reduce cross-correlation in error terms across stocks. The market return is CRSP value weighted average daily market return. In summary, our open–close return is defined as (current day’s close price/current day’s open price–1)–current day’s market return.

**Table 6.** Traders’ beliefs and stock returns.

	Open-Close Return	Close-Close Return	Close-Open Return
	(1)	(2)	(3)
Intercept	−0.0003 ** (0.0002)	−0.0005 *** (0.0001)	−0.0005 ** (0.0002)
MM. belief uncertainty	−0.0023 *** (0.0006)	−0.0014 ** (0.0006)	
Ifmd. belief uncertainty	0.0026 (0.0022)	−0.0011 (0.0015)	
Info. equality	−0.0003 *** (0.0001)	−0.0001 (0.0001)	
Order imbalance	0.0647 *** (0.0025)	0.0653 *** (0.0024)	
OI × MM. belief uncertainty	0.0582 *** (0.0070)	0.0544 *** (0.0079)	
Lag Return	0.0033 (0.0036)	−0.0083 ** (0.0035)	−0.0204 ** (0.0100)
Lag MM. belief uncertainty	−0.0004 (0.0005)	0.0018 *** (0.0005)	0.0070 *** (0.0024)
Lag Ifmd. belief uncertainty	0.0007 (0.0013)	−0.0006 (0.0015)	−0.0027 (0.0031)
Lag Info. equality	−0.0003 *** (0.0001)	−0.0001 (0.0001)	0.0006 *** (0.0002)
Lag OI	−0.0139 *** (0.0007)	−0.0164 *** (0.0007)	−0.0017 *** (0.0005)
R <sup>2</sup>	10.48%	9.96%	3.31%
Obs	893,176	893,160	893,179

This table reports results from Fama and MacBeth [26] regressions of stock returns on market makers belief uncertainty. MM. belief uncertainty is market makers’ belief uncertainty. Ifmd. belief uncertainty is informed traders’ belief uncertainty. Info. equality is informational equality. OI is order imbalance. See Appendix A for detailed definitions. The regressions are run on daily data. The dependent variables are stock returns. Column 1’s return is defined as (current day’s close price/current day’s open price–1)–current day’s market return. Column 2’s return is defined as (current day’s close price/previous day’s close price–1)–current day’s market return. Column 3’s return is defined as (current day’s open price/previous day’s close price–1). Market return is CRSP value weighted average daily market return. The open and close price are the midpoint quoted prices associated with the first and last transactions of the day, excluding the opening batch auction. Standard errors are in parentheses. \*, \*\*, and \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively.

Market makers’ belief uncertainty is negatively and significantly related to open–close return. This evidence supports the hypothesis that belief uncertainty depresses market makers’ willingness to hold the stock, leading to lower returns. Informed traders’ belief uncertainty does not have significant impact on the open–close return. This could be because informed traders do not dominate the trading. Therefore, the coefficient linked to informed traders’ belief uncertainty is insignificant. Informational equality is also negatively and significantly related to open–close return. It shows that after controlling for market makers’ belief uncertainty, the dominant effect of information asymmetry is to make price more informative. The regression also includes the previous day’s belief uncertainty and informational equality as explanatory variables. Only the coefficient of lag informational equality is significant, and it has the same sign as that of contemporaneous informational equality. We attribute this result to the persistence of informational equality.



Consistent with prior studies [34], the coefficient of contemporaneous order imbalance is positive, and that of the lag order imbalance is negative. Both coefficients are significant. Furthermore, the interaction term between order imbalance and market makers' belief uncertainty is positively and significantly associated with open–close return; this supports our model prediction that belief uncertainty causes market makers to adjust price more aggressively towards the direction of order imbalance.

In Column 2, we regress the conventional daily stock return with the same variables as those in Column 1. Column 2's return is defined as (current day's close price/previous day's close price–1)–current day's market return, where the prices are midpoint quoted prices. This regression partly serves as verification for the regression in Column 1. The coefficient of contemporaneous market makers' belief uncertainty is still negative and significant, but that of lag market makers' belief uncertainty becomes positive and significant. This result indicates that the previous day's uncertainty leads to a lower return on the previous day and a higher return on the current day. The difference in return definitions in Columns 1 and 2 could explain why the coefficient of lag market makers' belief uncertainty changes from Column 1 to 2. Column 2's return includes the price change from the previous day's close price to the current day's opening price. Thus, the coefficient changes probably because some information uncertainty is resolved after-hours. Accordingly, market makers are now more willing to hold the stock, causing the current day's open price to be higher relative to the previous day's depressed closing price. In other words, a major price rebound happens at the current day's market opening and this argument is supported by the next regression. In Column 2, the coefficients of both contemporaneous and lag informed traders' belief uncertainty and informational equality are all insignificant. Consistent with those in Column 1, the coefficients of order imbalance and its interaction term with market makers' belief uncertainty remain positive and significant, and the coefficient of lag order imbalance is still negative and significant.

To further explore the price rebound at the current day's market opening, the regression in Column 3 uses the price change from the previous day's close price to the current day's open price as a dependent variable. Here, the close–open return in Column 3 is defined as (current day's open price/previous day's close price–1). Consistent with the proposal that uncertainty resolution after-hours generates a price rebound at opening, both lag (previous day's) market makers' belief uncertainty and informational equality are positively and significantly associated with close–open return. Further supporting the price rebound explanation, the coefficients of lag (open–close) return and lag order imbalance are negative and significant.

Although not documented here, the same regressions were run on the samples excluding the period 2007–2009, when the recent financial crisis may have significantly impacted the market. The results are qualitatively similar with those reported in Table 6. The fact that belief parameters are persistent, including both contemporaneous and lag belief parameters in the regressions may have caused biased estimations in the regressions. We re-ran the regressions, excluding the lag belief parameters as explanatory variables, and the results were similar.

#### 4. Conclusions

Traders' beliefs are the key to understanding their trading behavior and further sustainable return. Considering the difficulty of directly observing traders' beliefs, this study develops a new approach to estimate traders' belief uncertainty and their relative informational advantages, which infers traders' belief parameters from their trading. The model divides investors into three categories: informed traders, market makers, and noise traders. The first two groups of traders follow Bayesian rules to update their beliefs after they observe order flows. Their newly updated beliefs lead to adjusted stock prices and new order flows. These belief–action interactions dictate the relation between stock price series and order flow series. As both data series are readily available, estimates of traders' beliefs can be obtained.



Traders' belief parameters provide more detailed pictures for market participants and can shed light on issues that are hard to answer with coarse belief proxies. Information asymmetry has long been considered an important factor in determining stock prices and sustainable returns; however, its effect has been vague in previous theoretical models. Here we illustrate that information asymmetry works in the interests of less informed traders, because it eventually makes price more informative. Furthermore, the uncertainty of less informed traders plays a much more important role than does information asymmetry.

These findings have several general implications. First, of the three major investor groups identified, it was the market makers who appeared to have most influence on trading and therefore market trends. Second, this important investor group utilizes diverse data, but the many parameters identified that impact on beliefs possibly mitigate bias due to asymmetry in information. Third, while stock trading is clearly influenced by a number of emotive factors, that are difficult to measure, the increased understanding of the diversity and relative importance of belief parameters will be helpful in prioritizing future research. By focusing further investigations on the most important belief parameters, or combinations, so as to reduce indecision, market volatility and enhance sustainable returns. In summary, despite acknowledged study limitations, any increase in predictability with lessened volatility contributes to the sustainability of the market—in terms of increased efficiency or reduced corrective trading—and so to sustainable returns.

**Author Contributions:** Conceptualization, Y.H., X.C. and X.-F.S.; methodology, X.C.; software, O.M.; validation, X.-G.Y. and K.J.B.; writing—original draft preparation, Y.H. and X.C.; writing—review and editing, X.-F.S. and X.-G.Y.; funding acquisition, Y.H.

**Funding:** This study was supported by the National Natural Science Foundation of China (Grant No. 71772013 and No. 71790604).

**Acknowledgments:** We deeply thank anonymous reviewers for their insightful suggestions and constructive comments. We are grateful to the editors for their patient work on our manuscript.

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

## Appendix A Variable Definitions

Variable Name	Definition
Time weighted average of percentage bid-ask spread	Percentage bid-ask spread is defined as $(P_a - P_b) / [0.5 \times (P_a + P_b)]$ , where $P_a$ and $P_b$ stand for the best bid and ask quote price, respectively. For each stock on each day, we compute the weighted average of the percentage spread, with the weights being the time this spread lasts.
1/Price	For each stock on each day, price is the volume weighted transaction prices during that day.
Volatility of transaction price	For each stock on each day, the standard deviation of transaction prices during that day.
log(size)	The logarithm of market capitalization of the firm on each day.
Turnover	For each stock on each day, turnover is defined as trading volume divided by total number of shares outstanding.
Order imbalance	For each stock on each day, order imbalance is defined as $(V_b - V_s) / (V_b + V_s)$ , where $V_b$ and $V_s$ stand for total volume of buy orders and sell orders, respectively. The buy orders and sell orders are determined according to Lee and Ready's (1991) method.

## References

1. Chung, K.H.; Elder, J.; Kim, J.-C. Corporate governance and liquidity. *J. Financ. Quant. Anal.* **2010**, *45*, 265–291. [\[CrossRef\]](#)
2. Alti, A.; Tetlock, P.C. Biased beliefs, asset prices, and investment: A structural approach. *J. Financ.* **2014**, *69*, 325–361. [\[CrossRef\]](#)
3. Kyle, A.S. Continuous auctions and insider trading. *Econometrica* **1985**, *53*, 1315–1335. [\[CrossRef\]](#)

4. Banerjee, S.; Green, B. Signal. or noise? uncertainty and learning about whether other traders are informed. *J. Financ. Econ.* **2015**, *117*, 398–423. [[CrossRef](#)]
5. Wang, J. A model of intertemporal asset prices under asymmetric information. *Rev. Econ. Stud.* **1993**, *60*, 249–282. [[CrossRef](#)]
6. Grossman, S.J.; Stiglitz, J.E. On the impossibility of informationally efficient Markets. *Am. Econ. Rev.* **1980**, *70*, 393–408.
7. Anderson, E.W.; Ghysels, E.; Juergens, J.L. Do heterogeneous beliefs matter for asset pricing? *Rev. Financ. Stud.* **2005**, *18*, 875–924. [[CrossRef](#)]
8. Banerjee, S. Learning from prices and the dispersion in beliefs. *Rev. Financ. Stud.* **2011**, *24*, 3025–3068. [[CrossRef](#)]
9. Hörner, J.; Lovo, S.; Tomala, T. Belief-free price formation. *J. Financ. Econ.* **2018**, *127*, 342–365. [[CrossRef](#)]
10. Vives, X. *Information and Learning in Markets: The Impact of Market Microstructure*; Princeton University Press: Princeton, NJ, USA, 2008.
11. Biais, B.; Foucault, T.; Moinas, S. Equilibrium fast trading. *J. Financ. Econ.* **2015**, *116*, 292–313. [[CrossRef](#)]
12. Easley, D.; de Prado, M.L.; O'Hara, M. Discerning information from trade data. *J. Financ. Econ.* **2016**, *120*, 269–285. [[CrossRef](#)]
13. Tannous, G.; Wang, J.; Wilson, C. The intraday pattern of information asymmetry, spread, and depth: Evidence from the NYSE. *Int. Rev. Financ.* **2013**, *13*, 215–240. [[CrossRef](#)]
14. Glosten, L.R.; Harris, L.E. Estimating the components of the bid/ask spread. *J. Financ. Econ.* **1988**, *21*, 123–142. [[CrossRef](#)]
15. Huang, R.D.; Stoll, H.R. The components of the bid-ask spread: A general approach. *Rev. Financ. Stud.* **1997**, *10*, 995–1034. [[CrossRef](#)]
16. George, T.J.; Kaul, G.; Nimalendran, M. Estimation of the bid-ask spread and its components: A new approach. *Rev. Financ. Stud.* **1991**, *4*, 623–656. [[CrossRef](#)]
17. Lin, J.-C.; Sanger, G.C.; Booth, G.G. Trade size and components of the bid-ask spread. *Rev. Financ. Stud.* **1995**, *8*, 1153–1183. [[CrossRef](#)]
18. Van Ness, B.F.; van Ness, R.A.; Warr, R.S. How well do adverse selection components measure adverse selection? *Financ. Manag.* **2001**, *30*, 77–98. [[CrossRef](#)]
19. Lee, C.M.; Ready, M.J. Inferring trade direction from intraday data. *J. Financ.* **1991**, *46*, 733–746. [[CrossRef](#)]
20. Bessembinder, H.; Venkataraman, K. Bid-ask spreads: Measuring trade execution costs in financial markets. *Encycl. Quant. Financ.* **2010**, 184–190.
21. Bessembinder, H. Issues in assessing trade execution costs. *J. Financ. Mark.* **2003**, *6*, 233–257. [[CrossRef](#)]
22. Boehmer, E.; Kelley, E.K. Institutional investors and the informational efficiency of prices. *Rev. Financ. Stud.* **2009**, *22*, 3563–3594. [[CrossRef](#)]
23. Brennan, M.J.; Chordia, T.; Subrahmanyam, A.; Tong, Q. Sell-order liquidity and the cross-section of expected stock returns. *J. Financ. Econ.* **2012**, *105*, 523–541. [[CrossRef](#)]
24. Bandi, F.M.; Russell, J.R. Microstructure noise, realized variance, and optimal sampling. *Rev. Econ. Stud.* **2008**, *75*, 339–369. [[CrossRef](#)]
25. O'Hara, M. High Frequency market microstructure. *J. Financ. Econ.* **2015**, *116*, 257–270. [[CrossRef](#)]
26. Fama, E.F.; MacBeth, J.D. Risk, return, and equilibrium: Empirical tests. *J. Political Econ.* **1973**, *81*, 607–636. [[CrossRef](#)]
27. Petersen, M.A. Estimating standard errors in finance panel data sets: Comparing approaches. *Rev. Financ. Stud.* **2009**, *22*, 435–480. [[CrossRef](#)]
28. McNish, T.H.; Wood, R.A. An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *J. Financ.* **1992**, *47*, 753–764. [[CrossRef](#)]
29. Chordia, T.; Sarkar, A.; Subrahmanyam, A. An empirical analysis of stock and bond market liquidity. *Rev. Financ. Stud.* **2004**, *18*, 85–129. [[CrossRef](#)]
30. Easley, D.; Hvidkjaer, S.; O'Hara, M. Is information risk a determinant of asset returns? *J. Financ.* **2002**, *57*, 2185–2221. [[CrossRef](#)]
31. Easley, D.; O'Hara, M. Information and the cost of capital. *J. Financ.* **2004**, *59*, 1553–1583. [[CrossRef](#)]
32. Lambert, R.A.; Leuz, C.; Verrecchia, R.E. Information asymmetry, information precision, and the cost of capital. *Rev. Financ.* **2011**, *16*, 1–29. [[CrossRef](#)]

33. Chordia, T.; Roll, R.; Subrahmanyam, A. Order imbalance, liquidity, and market returns. *J. Financ. Econ.* **2002**, *65*, 111–130. [[CrossRef](#)]
34. Chordia, T.; Subrahmanyam, A. Order imbalance and individual stock returns: Theory and evidence. *J. Financ. Econ.* **2004**, *72*, 485–518. [[CrossRef](#)]
35. Andrade, S.C.; Chang, C.; Seasholes, M.S. Trading imbalances, predictable reversals, and cross-stock price pressure. *J. Financ. Econ.* **2008**, *88*, 406–423. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).