Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Simple Multimarket Measure of Information Asymmetry

Travis L. Johnson, Eric C. So

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 64, No. 3, March 2018, pp. 1055–1080
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

A Simple Multimarket Measure of Information Asymmetry

Travis L. Johnson,[a] Eric C. So[b]

[a] McCombs School of Business, The University of Texas at Austin, Austin, Texas 78712; [b] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

Contact: travis.johnson@mccombs.utexas.edu (TLJ); eso@mit.edu (ECS)

Received: November 22, 2015
Revised: May 3, 2016
Accepted: July 1, 2016
Published Online in Articles in Advance: January 18, 2017

https://doi.org/10.1287/mnsc.2016.2608

Abstract. We develop and implement a new measure of information asymmetry among traders. Our measure is based on the intuition that informed traders are more likely than uninformed traders to generate abnormal volume in options or stock markets. We formalize this intuition theoretically and compute the resulting multimarket information asymmetry measure (MIA) for firm-days as a function of unsigned volume totals and without estimating a structural model. Empirically, MIA has many desirable properties: it is positively correlated with spreads, price impact, and absolute order imbalances; predicts future volatility; is an effective conditioning variable for trading strategies stemming from price pressure; and detects exogenous shocks to information asymmetry.

History: Accepted by Lauren Cohen, finance.

Keywords: information asymmetry • market microstructure • options • liquidity

1. Introduction

Asymmetric information plays a critical role in reconciling classical economic theory with observed economic behavior. The severity and content of asymmetric information influences most interactions between economic agents, particularly in cases of adverse selection or moral hazard. As a result, asymmetric information is one of the most fundamentally important concepts studied by modern economists. It is of particular interest to financial economists because one of the primary social benefits of an active securities market is that it aggregates disparate information into market prices. Therefore, understanding the influence of information asymmetry on the interaction between agents is essential for understanding the outcomes and operation of financial markets.

Most models of trading in financial markets involve two types of agents: those who trade because they have an information advantage and those who trade without an information advantage for reasons such as a desire for liquidity or hedging. The prevalence of informed trade affects liquidity, transaction costs, and trading volumes, and it can also help to explain market failures. For this reason, an important variable in most theoretical and empirical work on information economics in financial markets is the fraction of volume originating from informed traders.

Measuring the fraction of volume that is informed presents a significant challenge because it is inherently unobservable and time-varying. Thus, a central goal in studying information asymmetry in financial markets is to develop proxies that can be expressed as a function of observable inputs and whose variation captures empirically measurable repercussions of informed trade. In this paper, we address this measurement problem by developing a new proxy for information asymmetry among traders that leverages how trades are dispersed across option and stock markets. We empirically implement our measure and run a battery of tests that yield support for its use as a proxy for information asymmetry among traders.

Our multimarket information asymmetry measure, MIA, is a simple function of unsigned volume totals based on the idea that informed traders face a leverage constraint that generates a trade-off between smaller price impact in equity markets and additional leverage in options markets. Because informed traders receive correlated signals, this trade-off causes the fraction of informed trade occurring in options versus equity markets to fluctuate over time depending on the nature of the signals informed traders receive. In contrast, uninformed traders' choice of trading venues is relatively uncorrelated, and thus the fraction of uninformed trade in each market is relatively stable over time. As a result, periods of heightened information asymmetry manifest in abnormally high or low option-to-stock volume ratios (O/S, a measure first studied in Roll et al. 2010), relative to the level of O/S that occurs in the absence of private information. In Section 3, we formalize this intuition and derive MIA for firm j on day t as

$$\mathrm{MIA}_{j,t} = \frac{|O_{j,t}/S_{j,t} - M_{j,t}|}{O_{j,t}/S_{j,t} + M_{j,t}}, \tag{1}$$

where $M_{j,t}$ is an estimate of average O/S in the absence of informed trade.

Compared to alternative proxies for information asymmetry among traders, MIA has significant advantages in terms of its ease of implementation, clarity of interpretation, and empirical effectiveness. Alternative proxies generally fall within two broad categories: simple characteristics of the stock, and structural estimates of model parameters. Proxies based on simple characteristics like firm size and analyst coverage have the advantage of being easy to calculate and intuitively correlated with information asymmetry, but carry the disadvantages of being noisy and difficult to interpret because they are also correlated with several confounding factors such as liquidity and expected cash flows (e.g., in Lee and So 2017). Recognition of these disadvantages gave rise to structural estimates of information asymmetry among traders, including the "PIN" measure in Easley et al. (1996).[1]

Structural estimates such as PIN have the advantage of a theoretical connection to information asymmetry not confounded by liquidity or cash flows, but also carry the disadvantage of being computationally intensive because they require signing the direction of trades (i.e., buys versus sells), which has become increasingly problematic due to the accelerated frequency of trades in modern financial markets (Easley et al. 2012). Furthermore, Duarte et al. (2015) show that measures based on estimates of order imbalances alone, including PIN, are ineffective and note that "a different approach involving variables other than order flow is necessary to generate useful inferences about the arrival of informed trade" (p. 1).[2]

Our approach responds to this call by leveraging the dispersion of trades across options and stock markets. The addition of a second market allows us to identify the fraction of traders with an information advantage using abnormal volume imbalances across the two markets, as opposed to requiring estimates of imbalances in buyer- versus seller-initiated trades. Additionally, MIA can be estimated at the daily level, which allows researchers to study changes in information asymmetry in short windows surrounding information events.

We formalize the intuition behind MIA in a theoretical setting. As a baseline case, we first show that when a constant fraction of uninformed trade occurs in options and informed trading volume is concentrated entirely in either options or stock markets, MIA equals the fraction of volume originating from informed traders. We also assess MIA's effectiveness under more-realistic assumptions in a strategic trading model. Our model extends the Back (1993) framework to include a leverage constraint for informed traders that generates a trade-off between the price impact associated with concentrating volume in a single market and the extra leverage afforded by options. We also allow a random fraction of uninformed trading volume to occur in options markets, reflecting the possibility that uninformed trades, such as those emanating from mutual fund flows or index rebalancing, also generate variation in O/S.

In our extended model's equilibrium, although informed traders trade stocks and options simultaneously to mitigate price impact, they weigh the additional leverage and nonlinearity offered by options markets against the larger price impact, leading to substantial variations in the fraction of their volume they concentrate in options markets. As a result, we show deviations in O/S from its typical value indicate informed trade, and MIA is an effective proxy for information asymmetry, as long as the volatility of O/S driven by the random fluctuation of uninformed trader demands does not exceed the volatility of O/S driven by the informed trader's equilibrium strategy. In Section 2, we discuss empirical evidence from prior research suggesting this condition is met in practice.

Because our model shows MIA's effectiveness depends on the relative likelihood informed and uninformed traders generate abnormal O/S, we explore MIA's validity empirically. Specifically, we compute MIA for a panel of firm-days and subject it to a series of empirical tests. We divide our main empirical tests into three categories we refer to as (i) associations, (ii) predictions, and (iii) conditioning.

In our associations tests, we show MIA is positively associated with three repercussions of information asymmetry: bid-ask spreads, price impact, and absolute order imbalances. These results provide support for MIA and are consistent with the predictions of standard microstructure models (e.g., Glosten and Milgrom 1985, Kyle 1985) that illiquidity increases with the degree of information asymmetry among traders. In our prediction tests, we show that MIA positively predicts future volatility incremental to contemporaneous volumes and volatilities. This corroborates the predictions of microstucture models with time-varying information arrival (e.g., Easley and O'Hara 1987, Easley et al. 1998) that informed trade increases before the arrival of news. To mitigate concerns that MIA simply reflects expectations of volatility derived from public information, we also show that MIA predicts volatility incremental to option-implied volatility.

In our conditioning tests, we show that MIA helps distinguish between informed and uninformed sources of price pressure in equity and options markets. Specifically, we show that daily returns are less likely to reverse when MIA is high, consistent with the intuition in Llorente et al. (2002) that price changes driven by informed trade are less likely to reverse. We also show that the implied-volatility-spread trading strategy studied in Cremers and Weinbaum (2010) yields higher returns among firms with high MIA, consistent with implied-volatility spreads reflecting price pressure in options markets and MIA, indicating the extent to which the price pressure stems from informed trade. By showing that MIA, helps to explain variation in liquidity, volatility, and the returns to short-term trading strategies, our results illustrate several practical benefits of MIA as a measure of information asymmetry that can facilitate a host of interrelated asset allocation decisions.

In additional analyses, we show that MIA rises before, and declines immediately after, both earnings announcements and 8-K filings, consistent with MIA capturing short-window changes in information asymmetry around anticipated and unanticipated information events. Following Kelly and Ljungqvist (2012), we also identify exogenous shocks to information asymmetry driven by terminations of analyst coverage and show that MIA significantly increases following these shocks, whereas there are no significant changes in PIN.

An important contribution of our paper is to show that information asymmetry is better measured by absolute changes in O/S, rather than levels or signed changes of O/S. Roll et al. (2010), for example, suggest that O/S may be indicative of information asymmetry. Our model suggests that O/S alone is not a reliable proxy for the extent of information asymmetry because informed traders may prefer trading stocks over options, or vice versa, depending on the signal they receive and the relative leverage and transaction costs across the two markets. As a result, informed trading volume may be concentrated in either options or equity markets, meaning that both abnormally high and abnormally low O/S are indicative of informed trade. We empirically verify this result by showing that both increases and decreases in O/S predict future volatility, and that our main results are robust to controlling for the level of O/S.

Another related use of levels and signed changes in O/S are as proxies for the *sign* rather than *magnitude* of private information. Johnson and So (2012) shows that when there are short-sale costs in equity markets, O/S is a reliable proxy for the direction of private information. Specifically, increases in options volume indicate negative private information because informed traders use options more frequently when in possession of bad news to avoid the short-sale cost. In this paper, we focus on identifying the extent of private information among traders, which we show can be measured by absolute changes in O/S, rather than the sign of private information, which Johnson and So (2012) shows can be measured by levels or signed changes in O/S.

A possible alternative explanation for some of our empirical results is that volume in options markets rises prior to volatility events due to volatility or "vega" trading. However, several of our results indicate MIA detects the fraction of volume driven by private information about the mean asset value: the concentration of short-term reversal strategy returns among *low* MIA stocks; the concentration of implied-volatility-spread strategy returns among high MIA stocks; the ability of MIA to predict volatility incremental to implied volatility; and the ability of MIA to predict volatility even when O/S *decreases*. All of these results are consistent with MIA measuring privately informed directional trading but difficult to explain if MIA primarily reflects volatility trading.

2. Relation to Literature

Our model shows that the effectiveness of MIA depends on informed traders using sufficiently correlated strategies that they are more likely to generate abnormal O/S than uninformed traders. Prior research provides both theoretical and empirical support for this idea. On the theoretical front, Easley et al. (1996, 1997) model informed traders as each knowing that an asset is either overpriced or underpriced and therefore all trading in the same direction. On the empirical front, Cao et al. (2005) show that informed trade is more likely to occur in equity markets during normal times but is more likely to occur in options markets prior to extreme information events such as takeover announcements. Similarly, studies such as Anthony (1988), Stephan and Whaley (1990), and Chan et al. (2002) collectively show that information flows from option to equity markets, and vice versa, depending on the context. The combined evidence from these studies suggests that private information elicits fluctuations in how trade is allocated across options and stock markets, and thus that heightened information asymmetry manifests in abnormally high or low levels of O/S.

A natural point of comparison for MIA is PIN, the estimate of the probability of informed trade originating in Easley et al. (1996, 1997). Empirically estimating PIN requires a first stage estimate of the fraction of buyer- versus seller-initiated trades. The most common estimation technique is the algorithm developed in Lee and Ready (1991) that infers trade direction by comparing the trade price to the prevailing best bid and offer. However, Easley et al. (2012) point out that the increased execution speeds and evolving structure of modern financial markets have compromised the ability of the Lee–Ready algorithm to consistently infer the direction of trade from intraday data because quote and trade data are hard to merge accurately when both evolve at millisecond frequencies. Moreover, as Boehmer et al. (2007) and Asquith et al. (2010) discuss, measures of PIN inherit any shortcomings of trade classification algorithms used to calculate them. These facts together motivate us to develop a proxy for information asymmetry that does not require estimating the direction of trade and instead relies only on unsigned volume totals.

MIA has at least three important distinctions from PIN. The first is that instead of using imbalances within a market (between buys and sells), MIA is identified from volume imbalances across markets (between options and stocks). Volume imbalances across markets are easy to observe by comparing volume totals in the respective markets, circumventing the need to infer the direction of trade from intraday data. A second distinction of MIA relative to PIN is that our approach does not require estimating a structural model. Given a time series of order imbalance estimates, implementing PIN involves estimating the distributions of both informed and uniformed volumes using maximum likelihood applied to a longer time series of observations, traditionally done at the quarterly level. By contrast, implementing MIA does not involve estimating model parameters over a time series and thus can be estimated at shorter intervals, including the daily level. The third distinction is that MIA extends the scope of the analysis to measure informed trade across multiple markets. Prior research demonstrates that informed trade occurs in both options and equity markets, suggesting that our multimarket approach may yield a more comprehensive measure.

Odders-White and Ready (2008) provide an alternative structural approach to estimating the probability and magnitude of information events. Specifically, the authors use a modified Kyle (1985) model to summarize trading behavior and price setting each day, and estimate model parameters for firm-years using intraday data. Duarte et al. (2015) find that the Odders-White and Ready (2008) approach outperforms both the original PIN and the modified PIN in Duarte and Young (2009). Compared to MIA, the Odders-White and Ready (2008) approach has the advantage of separately identifying both the probability and magnitude of information asymmetry events, but has the disadvantages of computational complexity, a coarser annual frequency, and requiring first-stage estimates of order imbalances, which can be quite noisy in recent data.

3. Model and Empirical Predictions

3.1. Definition of MIA

We propose an empirically tractable proxy for multimarket information asymmetry, which we call MIA. The identifying intuition behind MIA is that informed traders face a leverage constraint that generates a tradeoff between smaller price impact in equity markets and additional leverage in options markets. Because informed traders receive correlated signals, this tradeoff causes the fraction of informed trade occurring in options versus equity markets to fluctuate over time depending on the nature of the information received. In contrast, uninformed traders' choice of trading venues is less correlated, and therefore the fraction of uninformed trade occurring in options markets is more stable over time. Combining these features implies that informed traders are more likely than uninformed traders to generate an abnormally high or low option-to-stock volume ratio O/S.

Based on this intuition, we define our multimarket information asymmetry measure, MIA, as is noted in Section 1:

$$\mathrm{MIA}_{j,t} = \frac{|O_{j,t}/S_{j,t} - M_{j,t}|}{O_{j,t}/S_{j,t} + M_{j,t}}, \tag{2}$$

where $O_{j,t}$ is the total option volume for firm j on day t, $S_{j,t}$ is the total equity volume, and $M_{j,t}$ is the median value of O/S for firm j. The numerator of Equation (2) captures the main intuition described above: Informed traders are more likely than uninformed traders to generate deviations in O/S from its typical levels, and so absolute differences between $O_{j,t}/S_{j,t}$ and $M_{j,t}$ are indicative of informed trade. The denominator of Equation (2) assures that MIA is nonnegative and converges to 1 as all volume becomes concentrated in either options or stock markets (O/S goes to infinity or zero).

3.2. Baseline Model: Ideal Conditions

To motivate our measure, we first consider a baseline case where MIA is an exact measure of information asymmetry. This occurs when we take our identifying assumption, that informed traders are more likely to generate abnormal cross-market volume imbalances than uninformed traders, to its extreme. To see this, consider a decomposition of stock and options trading volumes based on whether they originate from informed or uninformed traders:

$$S_t = S_{i,t} + S_{u,t}, \tag{3}$$

$$O_t = O_{i,t} + O_{u,t}. \tag{4}$$

Define u_t and i_t as the fraction of uninformed and informed trading volume occurring in options markets on day t:

$$u_t \equiv \frac{O_{u,t}}{O_{u,t} + S_{u,t}}, \tag{5}$$

$$i_t \equiv \frac{O_{i,t}}{O_{i,t} + S_{i,t}}. \tag{6}$$

Furthermore, the fraction of trades that are informed θ_t is given by

$$\theta_t \equiv \frac{O_{i,t} + S_{i,t} \cdot M}{O_t + S_t \cdot M}, \tag{7}$$

where M converts stock volume into an option volume representing the same number of trades, which we assume in our model and empirical tests equals O/S in the absence of informed trade. Using these definitions, we prove the following correspondence between MIA_t and θ_t when informed trade is concentrated entirely in either options or stock markets, while a constant fraction of uninformed trade occurs in options markets. These assumptions represent our identifying assumptions taken to an extreme because they imply *all* variation in O/S arises from informed trade, making MIA a precise measure of informed trade.

Theorem 1. *If informed traders use options or stock markets exclusively* ($i_t = \{0 \text{ or } 1\}$), *a constant fraction of uninformed traders use options* ($u_t = \bar{u}$), *and* $M = \bar{u}/(1-\bar{u})$, *then MIA exactly equals the fraction of trades that are informed* θ_t.

Proof. Under these assumptions, we only need to consider two possible outcomes, informed traders using options and informed traders using stock. In the former case, we have the following:

$$\text{MIA}_t \equiv \frac{|O_t - S_t \cdot M|}{O_t + S_t \cdot M} = \frac{|O_{i,t} + V_{u,t} \cdot \bar{u} - V_{u,t} \cdot (1-\bar{u}) \cdot M|}{O_t + S_t \cdot M} = \frac{O_{i,t}}{O_t + S_t \cdot M} = \theta_t, \tag{8}$$

where $V_{u,t} = O_{u,t} + S_{u,t}$. In the latter case, we have the following:

$$\text{MIA}_t \equiv \frac{|O_t - S_t \cdot M|}{O_t + S_t \cdot M} = \frac{|V_{u,t} \cdot \bar{u} - V_{u,t} \cdot (1-\bar{u}) \cdot M - S_{i,t} \cdot M|}{O_t + S_t \cdot M} = \frac{S_{i,t} \cdot M}{O_t + S_t \cdot M} = \theta_t. \tag{9}$$

Regardless of where informed traders concentrate their volume, we have $\text{MIA}_t = \theta_t$. □

Throughout our empirical analysis, we assume the average fraction of uninformed traders using options, M, is observable. We argue this is the case because there are a variety of settings in which there is very little informed trade, meaning the observed O/S indicates M. Our primary specification uses the firm's median O/S on recent trading days to estimate M under the assumption that the median impact of informed trade on O/S is zero. Under the parameterization of our model discussed in Section 3, we find the median O/S is extremely close to the true M. In Section 4.4, we also discuss using O/S in the days following earnings announcements or fitted values from a cross-sectional regression to estimate O/S in the absence of informed trade.

3.3. Extended Model: Strategic Trading

The assumptions in Theorem 1 are unlikely to hold in their exact form. For example, uninformed traders may generate unpredictable concentrations of volume in stock or options markets, making u_t different from the $\bar{u}$ used to calculate M. Similarly, option market makers may hedge in stock markets, or informed traders may receive different signals or seek to mitigate the price impact from trading in a single market, making i_t strictly between 0 and 1. To address these possibilities, we examine the robustness of MIA as a proxy for information asymmetry in a theoretical setting similar to Back (1993) where u_t fluctuates over time and i_t is often between 0 and 1.

Like Back (1993), our model extends the Kyle (1985) framework to allow strategic trading in options as well as the underlying stock. Because the informed trader takes into account their price impact, there is a natural incentive to use both options and stock markets simultaneously. The key addition we make to the Back (1993) model is a margin requirement, or equivalently a leverage constraint, limiting the position sizes of the informed trader. This requirement creates an interdependence of the informed trader's demand choice in options and stock markets, which is absent from the Back (1993) model because the informed trader is risk neutral. The margin requirement reflects the longstanding notion (e.g., in Black 1975, Easley et al. 1998) that informed traders prefer options relative to trading the stock directly because they offer additional leverage. Without a borrowing or margin limit, the additional leverage described in Black (1975) has no appeal to the informed trader.

3.3.1. Model Setup. A stock liquidates at $t = 1$ for $\tilde{v} \sim N(\bar{v}, \sigma_v^2)$, as do European call and put options with strike price $\bar{v}$ for:

$$\tilde{c} = (\tilde{v} - \bar{v})^+, \tag{10}$$
$$\tilde{p} = (\bar{v} - \tilde{v})^+, \tag{11}$$

respectively. The risk-free rate is zero and the stock does not pay dividends prior to $t = 1$. Trading occurs at time $t = 0$ between three types of agents: market makers, uninformed traders, and informed traders.

Uninformed traders' net demands for shares of stock $\tilde{z}_s$, calls $\tilde{z}_c$, and puts $\tilde{z}_p$ are independent and normally distributed with mean zero and variances $\sigma_{z,s}^2$, $\sigma_{z,c}^2$, and $\sigma_{z,p}^2$.

With probability ϕ, an informed trader observes $\tilde{v}$ at $t = 0$.[3] The informed trader chooses optimal demands $y_i \equiv (y_s, y_c, y_p)$ to maximize expected profits subject to the margin constraint:

$$m(y_i) \equiv |y_s| + \lambda(|y_c| + |y_p|) \le \bar{m}, \tag{12}$$

where $\bar{m}$ is their margin budget and λ is each option's margin use, both in units of shares. We assume $\lambda < 1$ to reflect the notion that options provide additional leverage relative to shares of stock. The informed trader's optimization given $\tilde{v} = v$ is therefore

$$y_i(v) = \underset{y \text{ s.t. } m(y)\le \bar{m}}{\arg\max}\ y_s(v - \mathbb{E}(\tilde{s}_0)) + y_c((v-\bar{v})^+ - \mathbb{E}(\tilde{c}_0)) + y_p((\bar{v}-v)^+ - \mathbb{E}(\tilde{p}_0)), \tag{13}$$

where $\tilde{s}_0$, $\tilde{c}_0$, and $\tilde{p}_0$ are the market clearing prices at time 0 for the three assets. These prices depend on the informed trader's choice of y and the uninformed traders' demand z. In choosing their demand $y_i(v)$, the informed trader takes into account the impact of their demand on expected prices. They compute these expected prices based on the equilibrium pricing functions market makers use and the distribution of possible uninformed trader demands.

Market makers observe total net order flow in each of the three markets, $x = y + z$, and set prices equal to the expected value of each asset given order flow in that market.[4] The pricing functions are therefore

$$s(x_s) = \mathbb{E}(\tilde{v} \mid y_s(\tilde{v}) + \tilde{z}_s = x_s), \tag{14}$$

$$c(x_c) = \mathbb{E}(\tilde{c} \mid y_c(\tilde{v}) + \tilde{z}_c = x_c), \tag{15}$$

$$p(x_p) = \mathbb{E}(\tilde{p} \mid y_p(\tilde{v}) + \tilde{z}_p = x_p). \tag{16}$$

3.3.2. Model Equilibrium. Equilibrium in the model consists of pricing functions $s(x_s)$, $c(x_c)$, and $p(x_p)$ based on conditional expectations given the equilibrium trading strategy of the informed trader $y(v)$, and trading strategy $y(v)$ that is optimal given the equilibrium pricing functions. The nonlinearity of both the options' payoffs and the margin constraint prevent us from deriving a closed-form solution. Instead, we solve

Table 1. Model Parameters and Simulations

Panel A: Parameter values

Parameter	Description	Value
$\bar{v}$	Average stock value	10
σ_v	Standard deviation of stock value	1
$\sigma_{z,s}$	Standard deviation of uninformed of stock demand	1
$\sigma_{z,c}$	Standard deviation of uninformed call demand	0.5
$\sigma_{z,p}$	Standard deviation of uninformed put demand	0.5
λ	Margin requirement for options relative to shares of stock	0.4
$\bar{m}$	Margin constraint in number of shares	1
ϕ	Probability there is an informed trader	25%, 50%, 75%, or 100%
σ_u	Standard deviation of fraction of uninformed volume in options	0%, 1%, 2%, or 5%
$\bar{u}$	Average fraction of uninformed volume in options	50%
V_u	Total volume of uninformed traders	3 or 6

Panel B: $\sigma_u = 0\%$

	$V_u = 3$				$V_u = 6$			
ϕ	0.25	0.50	0.75	1.00	0.25	0.50	0.75	1.00
$\hat{M}$	1.000	1.000	1.000	0.996	1.000	1.000	1.000	0.998
Mean θ (%)	5.14	9.92	14.08	17.62	2.87	2.87	7.77	9.67
Mean MIA (%)	1.16	2.13	3.12	3.89	0.66	1.19	1.72	2.12
$\sigma(O/S)$ inf (%)	3.51	3.29	3.01	2.69	2.00	1.85	1.66	1.47
$\sigma(O/S)$ no inf (%)	0.00	0.00	0.00	—	0.00	0.00	0.00	—
Mean MIA inf (%)	4.65	4.26	4.16	3.89	2.63	2.39	2.29	2.12
Mean MIA no inf (%)	0.00	0.00	0.00	—	0.00	0.00	0.00	—

Panel C: $\sigma_u = 1\%$

	$V_u = 3$				$V_u = 6$			
ϕ	0.25	0.50	0.75	1.00	0.25	0.50	0.75	1.00
$\hat{M}$	1.005	1.004	1.004	1.003	1.004	1.006	1.007	1.009
Mean θ (%)	5.15	9.92	14.08	17.62	2.87	5.51	7.77	9.66
Mean MIA (%)	2.46	3.15	3.76	4.18	2.02	2.32	2.55	2.66
$\sigma(O/S)$ inf (%)	3.59	3.39	3.11	2.82	2.18	2.05	1.88	1.72
$\sigma(O/S)$ no inf (%)	1.00	1.00	1.00	—	1.00	1.00	1.00	—
Mean MIA inf (%)	5.01	4.69	4.48	4.18	3.26	3.03	2.86	2.66
Mean MIA no inf (%)	1.61	1.60	1.60	—	1.61	1.61	1.62	—

Table 1. (Continued)

Panel D: $\sigma_u = 2\%$								
	$V_u = 3$				$V_u = 6$			
ϕ	0.25	0.50	0.75	1.00	0.25	0.50	0.75	1.00
$\hat{M}$	1.009	1.011	1.014	1.016	1.007	1.011	1.014	1.017
Mean θ (%)	5.14	9.91	14.06	17.60	2.87	5.51	7.76	9.66
Mean MIA (%)	3.85	4.32	4.69	4.88	3.47	3.63	3.70	3.69
$\sigma(O/S)$ inf (%)	3.85	3.66	3.42	3.16	2.67	2.57	2.44	2.33
$\sigma(O/S)$ no inf (%)	2.00	2.00	2.00	—	2.00	2.00	2.00	—
Mean MIA inf (%)	5.78	5.42	5.18	4.88	4.28	4.04	3.86	3.69
Mean MIA no inf (%)	3.21	3.22	3.24	—	3.20	3.22	3.25	—
Panel E: $\sigma_u = 5\%$								
	$V_u = 3$				$V_u = 6$			
ϕ	0.25	0.50	0.75	1.00	0.25	0.50	0.75	1.00
$\hat{M}$	1.016	1.024	1.031	1.034	1.011	1.017	1.021	1.022
Mean θ (%)	5.15	9.91	14.05	17.59	2.87	5.50	7.76	9.66
Mean MIA (%)	8.16	8.15	8.04	7.83	7.96	7.86	7.73	7.58
$\sigma(O/S)$ inf (%)	5.30	5.19	5.05	4.92	4.86	4.82	4.78	4.75
$\sigma(O/S)$ no inf (%)	5.00	5.00	5.00	—	5.00	5.00	5.00	—
Mean MIA inf (%)	8.60	8.26	8.03	7.83	7.86	7.72	7.63	7.58
Mean MIA no inf (%)	8.01	8.04	8.07	—	7.99	8.01	8.02	—

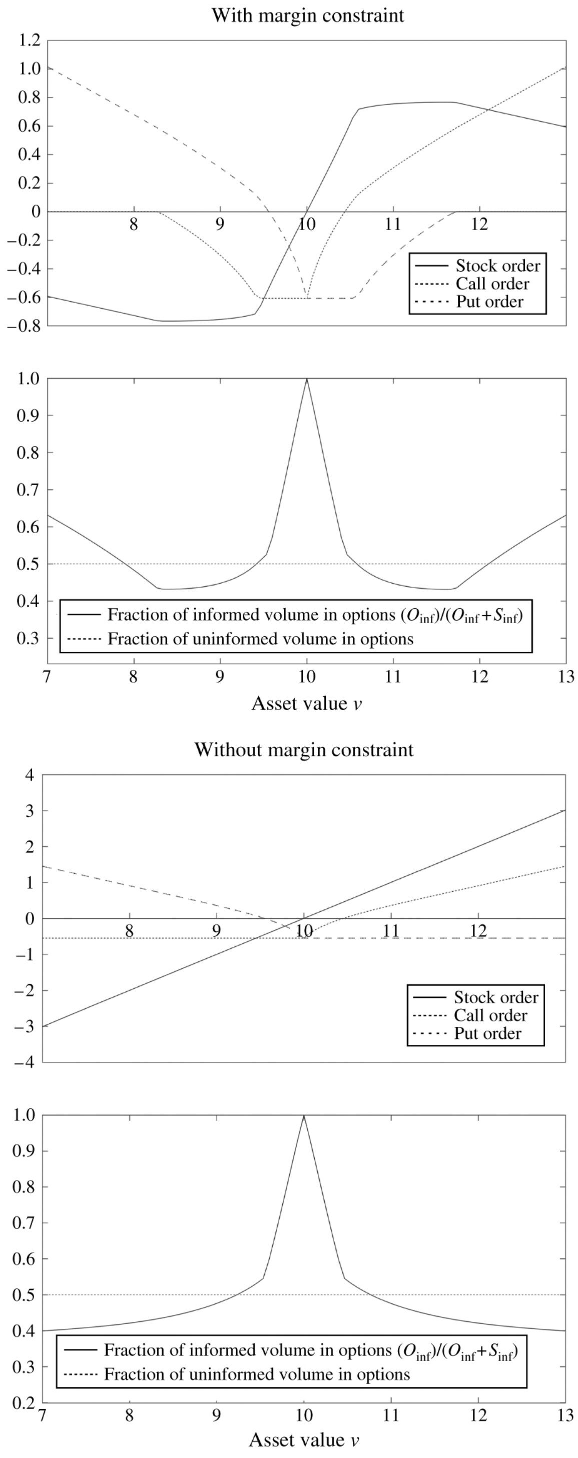
Notes. Panel A presents parameters used in numerical solutions of the model described in Section 3. When parameters have multiple listed values, we present results for each of the possibilities given. Panels B–E present results of 5,000,000 simulated observations based on numerical solutions of the model using different combinations of σ_u, θ, and V_u. For each simulation, we present $\hat{M}$, the estimated O/S in the absence of informed trade; average θ, the fraction of trades originating from informed traders; average MIA; the volatility of O/S and average MIA in observations where there is an informed trader; and the volatility of O/S and average MIA in observations without an informed trader.

the model numerically for a variety of parameter values given in panel A of Table 1.

The parameters we use are designed to be as realistic and generic as possible, and other reasonable parameter choices appear to produce qualitatively identical results. In Section 3.4, we discuss which parameter choices determine the effectiveness of MIA. The key addition we make to the Back (1993) model is a margin constraint, the impact of which is determined by λ, the margin requirement for options relative to shares of stock, and $\bar{m}$, the margin constraint. We assume options require a fraction $\lambda = 0.4$ of the margin required for stock positions, matching the ratio of the Regulation T margin requirements for naked short options positions (20% of the underlying stock price) and for short equity positions (50% of the underlying stock price). We set the margin constraint $\bar{m}$ equal to 1, meaning that if informed traders use stock markets exclusively they are limited to an order size equal to one standard deviation in uninformed order imbalances.

Figure 1 shows equilibrium demand functions of the informed trader given the parameters in Table 1. These demand choices reflect the tension between the desire to spread demand across assets to minimize price impact versus the nonlinearity and additional leverage afforded by options. When v is near its average value of 10, the informed trader cannot profitably trade the stock but can sell options, which will both expire with values near 0, less than their unconditional mean.[5] As v increases (decreases) from 10, the informed trader starts buying (selling) shares of stock and reducing their short positions in calls (puts). When v deviates from 10 by 0.5 or more, the informed trader's unconstrained portfolio choice violates the margin limit and so there is a "kink" in all three demand choices. As Figure 1 shows, the informed trader stops dramatically increasing their stock position for v beyond the kink, favoring options positions because of their smaller margin requirement. They also gradually phase out selling options as the asset value moves farther from its mean, eventually choosing $y_c = 0$ when v is sufficiently small, and $y_p = 0$ when v is sufficiently large. For even more extreme signals, the informed trader cuts back on their stock trades to free up margin capacity for even larger long positions in calls for good news and puts for bad news.

Figure 1. Equilibrium Informed Trader Demand in the Model



Notes. This figure presents the equilibrium trading strategy of informed traders in our model as a function of their information about the underlying asset's value v. The equilibrium is based on the parameters described in Table 1 with $\phi = 1$. In the first two panels, the margin constraint $\bar{m} = 1$. In the second two panels, there is no margin constraint. For each pair, the top panel presents their order sizes in shares of stock, call options, and put options. The bottom panel presents the fraction of their trading volume that occurs in options markets compared the fraction of uninformed trading volume.

The bottom plot in Figure 1 illustrates the key result of the model, that informed traders substantially vary the fraction of their volume occurring in options markets depending on the signal they receive. Their strategic concern for price impact causes them not to concentrate all their volume in one market, as required for the perfect correspondence in Section 3.2 to hold. However, the nonlinearity and additional leverage offered by options causes substantial variations in the fraction of informed trade in options markets. The fraction $\tilde{\imath}_t \equiv O_{i,t}/(O_{i,t} + S_{i,t})$ varies between 0.41 and 1. Moreover, this fraction is often substantially above or below the fraction of uninformed trade occurring in options, 0.5 in this parameterization. These fluctuations cause O/S to vary more substantially, which in turn increases average MIA, when there is more informed trade.

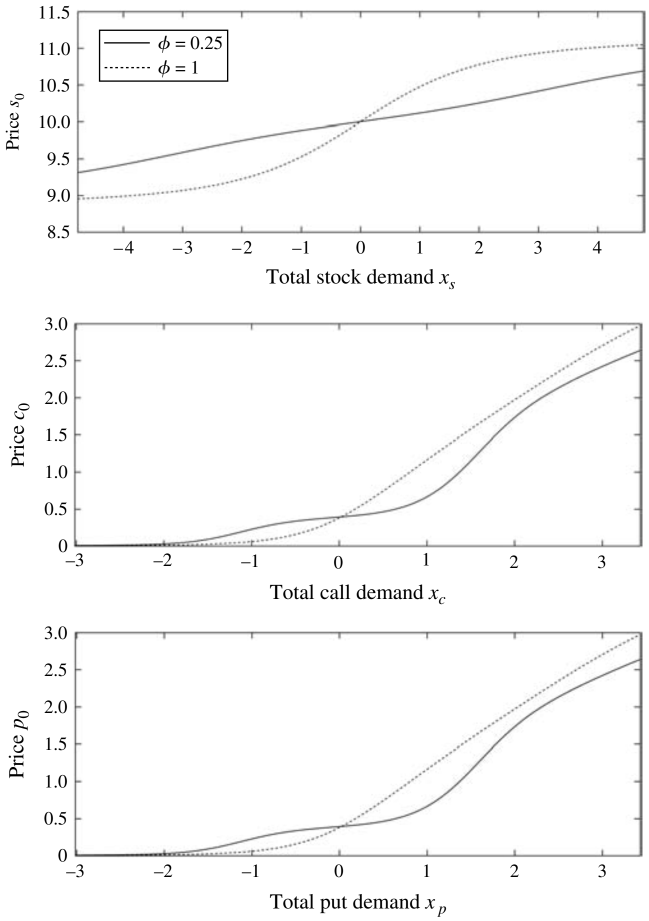
To illustrate the role of the margin constraint, the second pair of panels in Figure 1 show the equilibrium demand choices by informed traders in our model when $\bar{m} = \infty$. Without a margin constraint, the risk-neutral informed trader chooses volumes in each of the three markets independently. As a result, their demand in stock markets is the familiar linear function from Kyle (1985) and Back (1993). Their demand for both calls and puts follows the nonlinearity of the underlying assets, with the informed trader selling a fixed quantity of the option they know will expire out of the money, and buying an increasing amount of the option that will be in the money. For asset values far from 10, the equilibrium order sizes for options are smaller than the demands for stocks because there is more uninformed trade in stocks than calls or puts. As a result, the fourth panel of Figure 1 shows that the fraction of informed trade in options markets peaks for signals near zero and continually decreases as v gets farther from 10. Because the nonlinearity in options generates variations in informed O/S even without a margin constraint, MIA would still be an effective proxy for informed trade as long as uninformed O/S is sufficiently stable. However, we rely on the calibration with $\bar{m} = 1$ for our main analysis because it matches the empirical evidence that informed traders use options due to leverage constraints, and more importantly because it generates larger variations in informed O/S.

Figure 2 shows the equilibrium asset prices as a function of net order flow x. Price impact in equity markets is relatively small, while price impact in options markets is large and nonlinear, reflecting the convexity in option values as well as informed trader demands. Figure 2 also shows that price impacts are increasing in ϕ due to the additional risk of adverse selection.

3.3.3. Model Volume Totals.

To compute MIA in our model, we need expressions for the total volume occurring in options and stock markets. However, the Kyle (1985) framework does not naturally produce volume

Figure 2. Equilibrium Price Functions in the Model



Notes. This figure presents the equilibrium pricing function used by market makers as a function of the net demand in shares of stock, call options, and put options. The equilibrium is based on the parameters described in Table 1 with $\phi = 1$ or $\phi = 0.25$. Because the market makers are risk neutral and competitive, these prices equal the expected value of each asset conditional on the partially-informed order flow in that asset.

totals because it nets out demands among groups of traders and focuses on the equilibrium price as a function of the remaining order imbalances. We address this by separately modeling volume totals, which include offsetting trades as well as order imbalances, from the imbalances x and y specified by our model. Because the informed trader has no incentive to submit offsetting orders, we assume that when $y > 0$, they only submit buy orders; and when $y < 0$, they only submit sell orders, meaning informed trading volume equals $|y|$. The addition of multiple informed traders submitting potentially offsetting orders yields qualitatively identical results.

Our model allows uninformed traders, by contrast, to submit offsetting orders. Specifically, we model uninformed volume totals in options and equity markets independently as

$$O_u = V_u \cdot \tilde{u}, \tag{17}$$
$$S_u = V_u \cdot (1 - \tilde{u}). \tag{18}$$

Any such volume totals are consistent with uninformed order imbalances z_s, z_c, and z_p as long as the uninformed buy volume in options equals $(V_u \cdot \tilde{u} + z_p + z_c)/2$, uninformed sell volume in options equals $(V_u \cdot \tilde{u} - z_p - z_c)/2$, uninformed buy volume in stocks equals $(V_u \cdot (1 - \tilde{u}) + z_s)/2$, and uninformed sell volume in stocks equals $(V_u \cdot (1 - \tilde{u}) - z_s)/2$. Modeling volume totals separately from order imbalances allows us to study how changes in uninformed trading volume totals affect MIA while holding constant the equilibrium pricing function.

3.4. Simulations

A key result in our model, as illustrated by Figure 1, is that there is substantial variation in the fraction of informed trading volume occurring in options markets. We therefore hypothesize that abnormal O/S indicates informed trade, making MIA an effective measure, as long as the fraction of uninformed trading volume in options markets is less volatile than the fraction of informed volume. We test this hypothesis by simulating 5,000,000 observations based on numerical solutions for the demand and pricing functions given the parameterization of our model in Table 1.

For each simulated observation, we compute uninformed volume totals $O_{u,t}$ and $S_{u,t}$ using the approach outline in Section 3.3.3 and assuming $\tilde{u} \sim N(\bar{u}, \sigma_u^2)$ and $V_{u,t} = V_u$.[6] We also compute the informed volume totals $O_{i,t}$ and $S_{i,t}$ for each observation, which equal zero with probability $1 - \phi$ and are based on a randomly drawn $\tilde{v}$ and the equilibrium demand functions $y(v)$ with probability ϕ. We then combine these volume totals to compute both θ and MIA, as discussed in Section 3.2:

$$\theta_t \equiv \frac{O_{i,t} + S_{i,t} \cdot M}{O_t + S_t \cdot M} \qquad \mathrm{MIA}_t \equiv \frac{|O_t - S_t \cdot M|}{O_t + S_t \cdot M}. \tag{19}$$

To match our empirical implementation, we use the cross-observation median of O/S as M. Table 1 shows that across all values of σ_u, ϕ, and V_u, our estimate of M is extremely close to the true $M = \bar{u}/(1 - \bar{u}) = 1$ in our parameterization. The median O/S is an accurate estimate of M in our model because the median O/S observation either has no informed trading, which occurs with probability $1 - \theta$, or informed trading demand with O/S similar to the uninformed O/S.

Table 1 presents the results of these simulations for four different values of σ_u (0%, 1%, 2%, and 5%), four different values of ϕ (25%, 50%, 75%, and 100%), and two different values of V_u (3 and 6), yielding a total of 32 different simulated samples. We measure the effectiveness of MIA_t as a proxy for θ_t in two ways. The first is whether average MIA increases across simulated samples as we increase ϕ or decrease V_u, both of which should increase the prevalence of informed trade. The second is whether average MIA is higher within simulated samples among the observations where there is informed trade compared to the observations where there is no informed trade.

Panel B of Table 1 presents simulation results when $\sigma_u = 0$, meaning a constant fraction of uninformed volume occurs in options markets, as in the ideal conditions described in Section 3.2. As a result, all time variations in O/S are generated by informed traders, making MIA an excellent proxy for informed trade. Panel B shows that MIA (i) monotonically and proportionally increases in ϕ, (ii) is proportionally smaller when $V_u = 6$, and (iii) is exactly 0 in observations where there is no informed trade. The only shortcoming of MIA in this setting is that it understates θ_t. This is because informed traders use options and stock simultaneously, making the fraction of informed trade occurring in options i_t fluctuate between 0 and 1, rather than being equal to 0 or 1 as is required in Section 3.2. For this reason, and others discussed below, our empirical tests use MIA as an ordinal rather than cardinal measure of information asymmetry.

Panels C–E of Table 1 show simulation results as σ_u increases from 1% to 5%. They show that as long as O/S is more volatile across observations with informed trade than without informed trade, MIA is still an effective proxy for θ_t. For $\sigma_u = 1\%$ and $\sigma_u = 2\%$, we again find that MIA is increasing in ϕ, decreasing in V_u, and higher in observations with informed trade than those without informed trade. However, when $\sigma_u = 5\%$ the volatility of the fraction of uninformed trade occurring in options is very close to the volatility of the fraction of informed trade, making the volatility of O/S similar regardless of whether there is informed trade. In this case, MIA slightly decreases as ϕ increases, slightly decreases as V_u increases, and may be higher or lower in observations with informed trade. In untabulated results, we find that this trend continues and MIA is negatively related to θ for σ_u larger than 5%.[7] In practice, we expect σ_u to be smaller than the volatility of the fraction of informed trade occurring in options to the extent that uninformed traders receive less-correlated signals than informed traders.

Combined, the evidence in Figure 1 and Table 1 indicates MIA is an effective proxy for information asymmetry whenever the fraction of uninformed volume concentrated in options is sufficiently stable relative to the equilibrium fraction of informed demand in options markets. Our model illustrates how informed traders with leverage constraints will endogenously choose different volume concentrations in options markets depending on the signal they receive despite their desire to minimize price impact. However, since it is plausible for either informed or uninformed traders to be more likely to generate abnormal O/S, whether our identifying assumption holds and MIA is a good proxy for informed trade is ultimately an empirical question. To address this question, we estimate MIA for a panel of firm-days and subject it to a battery of tests designed to assess its validity as a proxy for information asymmetry among traders.

4. Empirical Tests

4.1. Sample Construction

The option data for this study come from the Ivy OptionMetrics database,[8] which provides end-of-day summary statistics for all exchange-listed options on U.S. equities. The summary statistics include option volume, implied volatilities, and option Greeks. The OptionMetrics database, and hence the sample for this study, spans from 1996 through 2013. We obtain daily equity trading volume, returns, and price data from the Center for Research in Security Prices (CRSP). We obtain quarterly financials from the Compustat Industrial Quarterly file. Our final sample is dictated by the intersection of OptionMetrics, Compustat, and CRSP data.

We restrict our sample to firm-days with recently active put and call options markets to ensure informed traders have access to options markets. Specifically, we require that each daily observation has a minimum of 30 trading days with positive volume in call and put contracts over the six months immediately preceding the measurement date in order to calibrate historical levels of firms' O/S ratios and to ensure traders have access to both markets. Requiring a recently active options markets, rather than a currently active options markets, is important in our setting because informed traders may trade in equity markets rather than option markets depending on the signal they receive. When a firm-day has a recently active options market, we set missing volume totals to zero when calculating MIA. We also eliminate closed-end funds, real estate investment trusts, American depository receipts, and firms with a stock price below \$1. The intersection of these databases and data restrictions results in 3,533,826 firm-days over 4,284 total trading days.

We calculate $\mathrm{MIA}_{j,t}$ using Equation (2):

$$\mathrm{MIA}_{j,t} = \frac{|O_{j,t}/S_{j,t} - M_{j,t}|}{O_{j,t}/S_{j,t} + M_{j,t}}, \tag{20}$$

where $O_{j,t}$ is the total option volume for firm j on day t and $S_{j,t}$ is the total stock volume. Specifically, following Johnson and So (2012), $O_{j,t}$ equals the total volume in option contracts, including both calls and puts, across all strikes for options expiring in the 60 trading days beginning five days after the trade date. We use daily stock volume $S_{j,t}$ from CRSP divided by a factor of 100 to make it more comparable to the quantity of option contracts that each pertain to 100 shares.

Table 2. Sample Descriptive Statistics

Panel A: Sample characteristics by year

	Firms	Firm-days	%CC	MEAN	MEDIAN	SD
1996	664	74,978	0.650	0.427	0.347	0.280
1997	883	127,236	0.763	0.438	0.354	0.283
1998	1,045	146,497	0.806	0.450	0.364	0.289
1999	1,134	160,268	0.860	0.430	0.350	0.282
2000	1,380	201,137	0.838	0.443	0.356	0.288
2001	1,254	178,232	0.738	0.494	0.405	0.299
2002	1,118	168,768	0.666	0.493	0.413	0.304
2003	1,047	158,754	0.865	0.469	0.385	0.290
2004	1,215	191,111	0.800	0.464	0.388	0.291
2005	1,316	203,525	0.764	0.460	0.385	0.287
2006	1,470	236,129	0.829	0.456	0.384	0.288
2007	1,645	257,755	0.885	0.468	0.395	0.290
2008	1,656	270,077	0.778	0.485	0.407	0.300
2009	1,501	234,094	0.926	0.497	0.399	0.300
2010	1,466	243,205	0.796	0.469	0.396	0.293
2011	1,471	242,739	0.890	0.459	0.378	0.295
2012	1,361	216,386	0.907	0.466	0.379	0.293
2013	1,368	222,935	0.930	0.453	0.368	0.288
All	1,307	201,302	0.814	0.465	0.384	0.292

Panel B: Descriptive statistics across MIA quintiles

	SIZE	LBM	COV	DISP	O	S	O/S
1 (Low MIA)	15.282	0.307	12.174	0.389	316,786	3,507,197	6.294
2	15.212	0.309	11.963	0.391	283,740	3,250,548	6.199
3	15.058	0.314	11.416	0.410	211,817	2,679,653	6.090
4	14.830	0.323	10.589	0.440	153,337	2,008,917	6.565
5 (High MIA)	14.338	0.350	8.979	0.759	112,715	1,603,517	7.034
High – Low	−0.943	0.043	−3.196	0.370	−204,070	−1,903,679	0.740

Panel C: Average correlations

	MIA	SIZE	LBM	COV	DISP	O/S
MIA		−0.228	0.075	−0.142	0.064	0.005
SIZE	−0.202		−0.094	0.471	−0.220	0.129
LBM	0.083	−0.093		−0.073	0.154	−0.104
COV	−0.122	0.438	−0.073		−0.185	0.119
DISP	0.064	−0.339	0.385	−0.121		−0.014
O/S	−0.260	0.189	−0.159	0.074	−0.054	

Notes. Panel A provides descriptive statistics of our multimarket information asymmetry measure, MIA. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. Panel A also lists the total number of firms and firm-days during our 1996–2013 sample. %CC is the fraction of market capitalization in the sample in June of each year relative to the total market capitalization of all U.S. equities (with share codes of 11 or 12) at the intersection of CRSP/Compustat. MEAN, MEDIAN, and SD refer, respectively, to the annual mean, median, and standard deviation of MIA. Panel B provides descriptive statistics across quintiles of MIA, formed each day. Panel C presents average daily Pearson (Spearman) correlations above (below) the diagonal. SIZE equals the log of market capitalization, and LBM equals the log of one plus the ratio of book-to-market equity. COV and DISP equal the number of analysts providing coverage and the dispersion in annual earnings forecasts measured 10 days prior to MIA in percent. O equals average daily option market volume multiplied by 100, S equals average daily stock market volume, and O/S is the percent option-to-stock volume ratio (winsorized daily at the 1% and 99% levels).

Calculating MIA also requires an estimate of $M_{j,t}$, the typical O/S value for firm j that would prevail in the absence of private information. We use the firm's median O/S over the past six months ending 10 trading days before t as our estimate of M, assuming the median O/S occurs on a day with no private information.[9] In Section 4.4, we show that our main results are robust in estimating M using either a cross-sectional approach, average levels of O/S following the firm's most recent earnings announcement, or when omitting option expiration weeks.

Panel A of Table 2 contains sample statistics for each year in our 1996–2013 sample window. The %CC column indicates that the firms in our sample account for roughly 81% of the market capitalization in the CRSP/Compustat universe as of the end of June for each calendar year. Additionally, %CC significantly increases over time from approximately 65% in 1996 to

93% in 2013, indicating that the continued growth of the options market increases the breadth of firms represented in our sample. As a point of comparison, our sample is similar in size and composition to the population of firms with analyst coverage. There are also exchanges for options on individual stocks in over 20 countries, opening the possibility of computing MIA in non-U.S. settings.

Panel A of Table 2 also reports descriptive statistics of MIA. The mean of MIA is fairly constant over time, ranging from a low of 0.427 in 1996 to a high of 0.497 in 2009. The full-sample average of MIA is 0.465, while the median is considerably lower at 0.384, suggesting that the distribution of MIA is right skewed. By construction, MIA is between 0 and 1, but there is significant cross-sectional variation, as indicated by the standard deviation of 0.292.

The average values of MIA in Table 2 may be larger than some readers find intuitive. Under the ideal conditions for MIA presented in Section 3.2, only informed traders generate abnormal O/S, and so MIA is the exact fraction of traders with private information about the firm's fundamental value. In reality, uninformed traders may also cause abnormal O/S in some cases, and informed traders may not always generate abnormal O/S, making an ordinal, rather than cardinal, measure of information asymmetry among traders. Our calibrated model reflects this, as the average values of MIA can be smaller or larger than the extent of informed trade θ depending on the parameterization, but there is nevertheless a strong correlation between MIA and θ.

Panel B of Table 2 provides time-series averages of firm characteristics across quintiles of MIA. We assign firms to quintiles each trading day, where the highest (lowest) values are assigned to quintile 5 (1). For each firm, we calculate SIZE, defined as the log of market capitalization; LBM, defined as the log of one plus the book-to-market ratio; COV, defined as the number of analysts covering the firm; and DISP, defined as the dispersion in analysts' FY1 earnings forecast. O/S is the daily option-to-stock volume ratio, winsorized at the 1% and 99% levels each trading day. Panel B demonstrates that SIZE, COV, and options and equity volumes (denoted as O and S, respectively) are all decreasing across MIA quintiles, which indicates that informed trade is more pronounced among smaller firms with less analyst coverage and lower trading volume. Similarly, LBM and DISP are increasing across MIA quintiles, indicating that informed trade is more pronounced among neglected "value" stocks with greater uncertainty.

Panel B of Table 2 also shows that O/S displays a U-shaped pattern across MIA quintiles. Similarly, panel C of Table 2 presents average daily correlations in firm characteristics and volume totals, which show that O/S has a weak positive Pearson correlation (=0.005) with MIA but a negative Spearman correlation with MIA (= −0.260). This is consistent with the key result in our model that informed traders concentrate different fractions of their volume in options markets depending on the signal they receive, resulting in either an increase or a decrease in O/S (meaning informed trade is better measured by the absolute change in O/S than the level of O/S).

4.2. Associations with MIA

Most models of trading with asymmetric information share the feature that illiquidity, as measured by bid-ask spreads or price impact, increases in response to adverse selection risk. Because abnormal O/S, by definition, is difficult to anticipate, we predict illiquidity increases in both options and equity markets when adverse selection is high regardless of whether abnormal volume is currently concentrated in options or equity markets. Therefore, if MIA is an effective proxy for information asymmetry among traders, it should be positively related to bid-ask spreads and price impact regardless of where abnormal volume is currently concentrated.

We test this prediction first using relative bid-ask spreads in both equity and options markets. For equity markets, we calculate a firm's relative bid-ask spread at the end of each trading day using closing quotes from CRSP.[10] For options markets, we calculate the average end-of-day bid-ask spread for all outstanding options contracts for a given firm at the end of each trading day, weighting each option contract by its open interest.

Columns (1) and (2) of Table 3 contain results from daily Fama and MacBeth (1973) regressions of equity market spreads on contemporaneously measured MIA and several control variables.[11] The MIA coefficient in column (1) is 0.010 (t-statistic = 9.80), indicating that MIA possesses a significant positive relation with contemporaneous equity market spreads. This relation holds after controlling for the log of one plus contemporaneous changes in equity and option volumes, ΔEQVOL and ΔOPVOL, defined, respectively, as the percentage change in equity and option volumes relative to the average of each volume over the six months ending 10 days prior to day t. We also control for relative spreads from $t-1$ to $t-5$ to account for autocorrelation in spreads. Column (2) demonstrates that the positive MIA-spread relation is robust in controlling for O/S, consistent with the earlier evidence that MIA and O/S capture distinct market outcomes. By controlling for ΔEQVOL and ΔOPVOL, we absorb any impact that changes of O/S have on spreads, meaning the association between MIA and spreads is driven by the nonlinear functional form used to compute MIA rather than a simple linear relation between O/S and spreads.

Table 3. Relative Spreads and Illiquidity Regressions

Dep. variable:	Equity spreads		Option spreads		ILLIQ	
	(1)	(2)	(3)	(4)	(5)	(6)
MIA	0.010***	0.010***	0.002***	0.002***	0.003***	0.002***
	(9.80)	(9.92)	(7.07)	(7.21)	(3.74)	(3.18)
Lag(−1)	0.191***	0.191***	0.412***	0.412***	0.226***	0.225***
	(31.76)	(31.67)	(91.53)	(91.53)	(107.81)	(107.28)
Lag(−2)	0.173***	0.173***	0.184***	0.184***	0.202***	0.202***
	(34.62)	(34.52)	(68.44)	(68.34)	(102.30)	(101.43)
Lag(−3)	0.159***	0.159***	0.186***	0.186***	0.189***	0.188***
	(33.95)	(33.78)	(77.10)	(77.09)	(92.40)	(92.20)
Lag(−4)	0.162***	0.162***	0.202***	0.202***	0.182***	0.182***
	(35.14)	(35.19)	(75.01)	(75.08)	(87.33)	(87.06)
VLTY	0.000***	0.000***	−0.000***	−0.000***	0.000***	0.000***
	(2.62)	(2.79)	(−5.25)	(−5.26)	(10.84)	(9.39)
SIZE	−0.007***	−0.007***	−0.000	−0.000	−0.021***	−0.021***
	(−11.21)	(−10.96)	(−0.60)	(−0.45)	(−21.11)	(−21.01)
INST	−0.013***	−0.013***	0.002***	0.002***	−0.031***	−0.030***
	(−10.02)	(−10.01)	(5.49)	(5.63)	(−17.37)	(−17.69)
COV	0.001**	0.001**	−0.001***	−0.001***	0.003***	0.003***
	(2.33)	(2.23)	(−6.28)	(−6.42)	(9.02)	(8.66)
ΔEQVOL	−0.001*	−0.001**	0.001***	0.001***	−0.014***	−0.013***
	(−1.91)	(−2.02)	(15.34)	(15.30)	(−18.85)	(−19.20)
ΔOPVOL	−0.000*	−0.000	0.000	−0.000	0.000***	0.000
	(−1.85)	(−1.07)	(0.18)	(−0.21)	(7.08)	(1.14)
O/S	—	−0.000***	—	−0.000	—	0.000***
	—	(−4.49)	—	(−0.01)	—	(10.43)
Intercept	0.148***	0.146***	0.003**	0.003**	0.349***	0.353***
	(12.61)	(12.47)	(2.11)	(1.96)	(21.54)	(21.54)
R^2 (%)	46.477	46.575	58.571	58.660	64.233	64.352

Notes. This table presents average regression statistics from daily Fama–MacBeth regressions where the dependent variable is a firm's relative spread in equity markets; relative spread in options markets; or Amihud illiquidity ratio measured on day t, where t is the date of measurement for our multimarket information asymmetry measure, MIA. We calculate daily relative spreads in equity markets using closing bid and ask prices from CRSP. Option market spreads are defined as the open-interest-weighted average relative spreads across all option contracts for a given firm-day. Both spread measures are calculated on the same day as MIA. We calculate daily Amihud illiquity ratio (ILLIQ) as the daily absolute return scaled by total dollar volume and multiplied by 10^9. Lag(N) equals the dependent variable measured on day $t+N$. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. SIZE equals the log of market capitalization and VLTY is the standard deviation of returns over the six months ending 10 days prior to day t. INST is the fraction of shares held by institutions, and COV is the log of one plus the number of analysts covering the firm. ΔEQVOL and ΔOPVOL equal the percentage change in equity and option volumes relative to the average of each volume over the six months ending 10 days prior to day t. O/S is the option-to-stock volume ratio on day t. The parentheses contain t-statistics from the Fama–MacBeth regressions after Newey and West (1987) adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Columns (3) and (4) of Table 3 contain regression results where the dependent variable is option market spreads, also measured contemporaneously with MIA. The MIA coefficients in columns (3) and (4) are again significantly positive (t-statistics = 7.07 and 7.21, respectively), indicating MIA is associated with wider spreads in options markets. We also examine the association between MIA and an alternative measure of liquidity, the log Amihud illiquidity ratio (ILLIQ). Columns (5) and (6) of Table 3 show MIA is positively related to concurrent values of ILLIQ (t-statistics = 3.74 and 3.18, respectively), indicating there is greater price impact and less liquidity when MIA is high. Together, the evidence in Table 3 provides support for MIA as a measure of informed trade by showing that market makers decrease liquidity when adverse selection, as measured by MIA, is high.

Another implication of most models of trading under information asymmetry is that informed traders are more likely to generate order imbalances than uninformed traders. Therefore, if MIA is a good proxy for information asymmetry, it should be positively associated with order imbalances. To test this prediction, Table 4 examines the association between MIA and absolute order imbalances in the equity market estimated from daily trade and quote data using the Lee and Ready (1991) algorithm. Despite the aforementioned problems with merging trade and quote data

Table 4. Absolute Order Imbalance Regressions

Dep. variable:	OIB		
	(1)	(2)	(3)
MIA	0.015***	0.007***	0.006***
	(27.45)	(10.85)	(10.47)
MIA × D(O/S < M)	—	0.014***	0.014***
	—	(19.15)	(18.64)
D(O/S < M)	—	−0.002***	−0.002***
	—	(−5.00)	(−5.24)
OIB(−1)	0.128***	0.128***	0.127***
	(46.93)	(46.47)	(46.38)
OIB(−2)	0.096***	0.095***	0.095***
	(42.79)	(42.61)	(42.34)
OIB(−3)	0.088***	0.087***	0.087***
	(42.02)	(41.65)	(41.53)
OIB(−4)	0.085***	0.084***	0.084***
	(39.84)	(39.53)	(39.38)
VLTY	−0.001***	0.001***	0.001***
	(−21.03)	(5.80)	(5.89)
SIZE	−0.007***	−0.000	−0.000*
	(−26.25)	(−1.59)	(−1.92)
INST	−0.001**	−0.001***	−0.001***
	(−2.38)	(−20.91)	(−20.94)
COV	−0.004***	−0.007***	−0.007***
	(−19.88)	(−25.99)	(−25.10)
ΔEQVOL	0.001***	−0.004***	−0.004***
	(6.98)	(−20.18)	(−20.10)
ΔOPVOL	−0.000***	−0.001**	−0.001*
	(−8.18)	(−2.00)	(−1.66)
O/S	—	—	0.000
	—	—	(0.48)
Intercept	0.198***	0.198***	0.198***
	(33.89)	(33.71)	(33.09)
R^2 (%)	13.515	13.895	14.093

Notes. This table presents average regression statistics from daily Fama–MacBeth regressions where the dependent variable is a firm's absolute order imbalance measured on day t, where t indicates the date of measurement for our multimarket information asymmetry measure, MIA. We calculate daily absolute order imbalance as the absolute imbalance of buyer- versus seller-initiated volume scaled by total volume. We identify buyer- versus seller-initiated volume using the Lee–Ready algorithm and intraday TAQ data. OIB(N) equals the absolute order imbalance on day $t+N$. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. SIZE equals the log of market capitalization and VLTY is the standard deviation of returns over the six months ending 10 days prior to day t. INST is the fraction of shares held by institutions, and COV is the log of one plus the number of analysts covering the firm. ΔEQVOL and ΔOPVOL equal the percentage change in equity and option volumes relative to the average of each volume over the six months ending 10 days prior to day t. (O/S < M) is an indicator variable equal to one when O/S is less than M, where M is the median level of O/S over the past six months ending 10 days prior. The parentheses contain t-statistics from the Fama–MacBeth regressions after Newey–West adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

in high-frequency markets, estimates of order imbalances are still commonly used by economists studying information in financial markets; thus, establishing a relationship between MIA and order imbalances is a helpful validity check for our measure. Table 4 shows that higher values of MIA are associated with greater absolute order imbalances, incremental to past imbalances and concurrent changes in volumes. This relation is consistent with the intuition that informed traders are more likely to create directional order imbalances as they attempt to capitalize on their information.

Our model predicts that equity market imbalances are more closely related to the extent of information asymmetry when informed traders concentrate a higher fraction of their volume in equity markets, making O/S smaller than M. To test this prediction, we add an interaction term between MIA and an indicator variable that equals one when O/S < M. Consistent with our prediction, we find that the relation between MIA and equity market order imbalances is twice as strong when O/S < M, compared to cases when O/S > M. Note that we do not include interactions between MIA and O/S < M in our spread and price-impact regressions (i.e., Table 3) based on the intuition that equity market makers are concerned about adverse selection risk when $\text{MIA}_{j,t}$ is high, and therefore reduce liquidity, without knowing ahead of time where informed agents actually trade on day t.

4.3. Predicting Volatility with MIA

Our next tests are based on the hypothesis that informed trade is more prevalent before periods of abnormal volatility. This could be the case for three reasons. The first is that some traders, through the use of prediction models or privileged access to the information itself, may have foreknowledge of pending news, resulting in an increase in the number of informed traders and abnormally high future volatility when the news becomes public (e.g., Grossman and Stiglitz 1980). The second is that, controlling for the number of informed traders, higher volatility news presents a more profitable trading opportunity and will therefore increase the size of informed traders' orders (e.g., Easley and O'Hara 1987). The third is that informed traders themselves create an order imbalance cascade that causes the subsequent volatility in stock prices (e.g., Easley et al. 2012). All of these forces suggest that next-period returns will be more volatile when MIA is high, both when abnormal volume is concentrated in options and when it is concentrated in stocks.

We test this prediction by examining the relation between MIA and future realized volatility in a multivariate setting. Column (1) of panel A in Table 5 contains results from Fama–MacBeth regressions of one-day-ahead squared returns, RETSQ, on

Table 5. Predicting Volatility Using MIA

Panel A: Regressions of future volatility on MIA					
Dep. variable:	RETSQ(1) (1)	RETSQ(2) (2)	RETSQ(3) (3)	RETSQ(4) (4)	RETSQ(5) (5)
MIA	1.307***	0.815***	0.994***	1.023***	0.671**
	(5.04)	(3.18)	(3.38)	(3.63)	(2.40)
RETSQ	0.132***	0.083***	0.079***	0.051***	0.062***
	(17.64)	(9.26)	(11.01)	(8.47)	(6.83)
RETSQ(−1)	0.070***	0.073***	0.052***	0.060***	0.067***
	(9.11)	(10.71)	(9.00)	(7.13)	(7.21)
RETSQ(−2)	0.062***	0.055***	0.065***	0.074***	0.061***
	(10.47)	(10.19)	(7.98)	(7.73)	(7.59)
RETSQ(−3)	0.047***	0.064***	0.068***	0.067***	0.044***
	(9.38)	(7.54)	(7.28)	(7.97)	(7.66)
RETSQ(−4)	0.062***	0.067***	0.065***	0.042***	0.057***
	(8.14)	(7.25)	(8.81)	(6.45)	(10.26)
VLTY	0.483***	0.501***	0.513***	0.538***	0.545***
	(22.53)	(21.70)	(21.35)	(22.18)	(22.31)
SIZE	−0.928***	−0.977***	−0.991***	−1.011***	−1.022***
	(−11.23)	(−12.67)	(−12.66)	(−12.66)	(−12.43)
INST	−1.291***	−1.228***	−1.297***	−1.387***	−1.382***
	(−4.05)	(−3.59)	(−3.57)	(−3.86)	(−3.87)
COV	0.327***	0.263***	0.246***	0.221***	0.211**
	(4.15)	(3.36)	(3.08)	(2.69)	(2.53)
ΔEQVOL	1.730***	0.798***	0.496***	0.573***	0.299***
	(9.68)	(5.66)	(4.60)	(4.81)	(2.64)
ΔOPVOL	−0.021**	−0.031***	−0.029***	−0.044***	−0.041***
	(−2.33)	(−3.22)	(−2.96)	(−3.52)	(−3.58)
O/S	0.057***	0.060***	0.070***	0.061***	0.065***
	(5.97)	(5.93)	(7.21)	(7.24)	(7.24)
Intercept	14.609***	16.229***	16.565***	17.155***	17.526***
	(11.08)	(12.22)	(12.27)	(12.39)	(12.41)
R^2 (%)	11.192	10.023	9.624	9.405	9.296

MIA and several control variables. We use Fama–MacBeth regressions to focus on cross-sectional variation in MIA rather than time-series variation that could be driven by aggregate factors such as changing market structure, macroeconomic conditions, trader sentiment, or funding liquidity. We also control for cross-sectional variation in lagged volatility over the past five trading days, longer-window historical volatility, firm size, O/S, and changes in options and equity volumes.

Column (1) of Table 5 shows that MIA positively predicts next-day volatility incremental to our control variables. As an indication of this relation's economic significance, all else equal, a one-standard-deviation increase in MIA increases next-day squared returns by 0.38 percentage points, corresponding to a 0.62 percentage point increase in volatility.[12] Columns (2)–(5) show that MIA has significant, gradually-diminishing, predictive power for future volatility from $t+2$ to $t+5$, suggesting that MIA captures information asymmetry about events that are likely to occur within the next week of trading.

Given the strong relation between MIA and future volatility documented in Table 5, one potential concern is that these results reflect a mechanical correlation between option volume and expectations of volatility based on public information. We document at least two results that mitigate this concern. First, in panel B of Table 5, we disaggregate MIA and show that it predicts volatility both when O/S increases (i.e., $O/S > M$) and when it decreases (i.e., $O/S < M$), which is consistent with our model because informed traders may prefer to trade stocks over options before the arrival of public information. Second, as discussed below, we also control for option-implied volatility as a summary measure of expected volatility based on public information.

To assess how much of the information in MIA about future volatility is already reflected in option prices and commonly used proxies for information asymmetry, we repeat our volatility prediction regressions with four additional independent variables. The first is option-implied variance, IV. We calculate IV concurrently with MIA on day t as the square of average implied volatility across the firm's closest maturity,

Table 5. (Continued)

Panel B: Regressions of future volatility on disaggregated MIA

Volatility measure:	RETSQ(1) (1)	RETSQ(2) (2)	RETSQ(3) (3)	RETSQ(4) (4)	RETSQ(5) (5)
MIA (O/S > M)	1.613***	0.696***	0.423	0.727**	−0.290
	(4.88)	(2.66)	(1.30)	(2.33)	(−0.99)
MIA (O/S < M)	1.031***	0.714**	1.183***	1.169***	1.083***
	(3.52)	(2.38)	(3.42)	(3.51)	(3.25)
RETSQ	0.131***	0.082***	0.079***	0.050***	0.062***
	(17.45)	(9.21)	(10.96)	(8.24)	(6.82)
RETSQ(−1)	0.069***	0.073***	0.051***	0.060***	0.067***
	(8.98)	(10.67)	(8.93)	(7.09)	(7.22)
RETSQ(−2)	0.062***	0.054***	0.064***	0.074***	0.060***
	(10.41)	(10.14)	(7.93)	(7.71)	(7.61)
RETSQ(−3)	0.046***	0.064***	0.068***	0.067***	0.044***
	(9.27)	(7.52)	(7.26)	(7.96)	(7.66)
RETSQ(−4)	0.061***	0.066***	0.065***	0.042***	0.057***
	(8.14)	(7.22)	(8.81)	(6.41)	(10.24)
VLTY	0.486***	0.502***	0.512***	0.538***	0.543***
	(22.85)	(21.84)	(21.48)	(22.37)	(22.37)
SIZE	−0.925***	−0.979***	−0.995***	−1.008***	−1.023***
	(−11.20)	(−12.80)	(−12.71)	(−12.65)	(−12.46)
INST	−1.288***	−1.213***	−1.252***	−1.346***	−1.295***
	(−4.12)	(−3.61)	(−3.55)	(−3.85)	(−3.71)
COV	0.324***	0.259***	0.237***	0.214***	0.197**
	(4.21)	(3.36)	(3.02)	(2.64)	(2.41)
ΔEQVOL	1.747***	0.811***	0.496***	0.570***	0.282**
	(9.69)	(5.75)	(4.61)	(4.74)	(2.47)
ΔOPVOL	−0.032***	−0.037***	−0.028**	−0.042***	−0.031**
	(−3.20)	(−3.47)	(−2.54)	(−2.69)	(−2.37)
O/S	0.057***	0.064***	0.081***	0.067***	0.080***
	(5.64)	(6.27)	(7.73)	(7.80)	(8.29)
Intercept	14.580***	16.282***	16.646***	17.131***	17.593***
	(11.04)	(12.32)	(12.32)	(12.36)	(12.46)
R^2 (%)	11.369	10.186	9.786	9.568	9.453

Notes. This table presents summary statistics from daily Fama–MacBeth regressions where the dependent variable is squared realized percent returns. RETSQ(N) equals the square of firms' raw percent return on day $t+N$, where t indicates the date of measurement for our multimarket information asymmetry measure, MIA. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. SIZE equals the log of market capitalization and VLTY is the standard deviation of returns over the six months ending 10 days prior to day t. INST is the fraction of shares held by institutions and COV is the log of one plus the number of analysts covering the firm. ΔEQVOL and ΔOPVOL equal the percentage change in equity and option volumes relative to the average of each volume over the six months ending 10 days prior to day t. Panel B disaggregates MIA based on whether O/S is larger than M, where M is the median level of O/S over the past six months ending 10 days prior. The parentheses contain t-statistics from the Fama–MacBeth regressions after Newey–West adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

at-the-money call and put options, divided by 252 to transform it into a daily measure. The next three are often-used proxies for information asymmetry among traders: the firm's relative spread, Amhiud illiquidity ratio, and absolute order imbalance, all measured concurrently with MIA.

The results of this horse race are in Table 6. The IV coefficient is positive, highly significant, and absorbs most of the predictive power of past volatility. More importantly, MIA is significantly positively related to future volatilities from days $t+1$ to $t+5$ (t-statistics ranging from 4.77 to 1.88) incremental to IV as well as the other controls and proxies in Table 4. The significantly positive relation is consistent with MIA capturing private information about future price movements that is not yet reflected in option prices or other proxies for information asymmetry. However, the coefficients on MIA in panel A of Table 6 are much smaller than the analogous results in Table 5 that do not control for IV, indicating that options markets reflect much, but not all, of the information in MIA about future volatility.[13] Finally, while MIA predicts volatility incremental to

Table 6. Incremental Volatility Prediction

Dep. variable:	RETSQ(1) (1)	RETSQ(2) (2)	RETSQ(3) (3)	RETSQ(4) (4)	RETSQ(5) (5)
MIA	1.065***	0.595***	0.781***	0.703***	0.465*
	(4.77)	(2.76)	(2.87)	(3.01)	(1.88)
RS	1.281	0.365	−1.138	−0.349	−0.450
	(1.18)	(0.33)	(−1.23)	(−0.26)	(−0.41)
ILLIQ	−1.808***	−1.498***	−1.181***	−1.599***	−1.119***
	(−4.33)	(−3.59)	(−2.61)	(−4.05)	(−2.61)
OIB	−3.196***	−2.221***	−1.490***	−0.903	−1.828***
	(−4.22)	(−4.55)	(−2.98)	(−1.21)	(−2.81)
O/S	−0.002	−0.003	0.003	0.003	0.004
	(−0.22)	(−0.33)	(0.36)	(0.36)	(0.41)
IV	0.878***	0.901***	0.932***	0.923***	0.857***
	(20.81)	(17.29)	(18.09)	(16.94)	(18.13)
RETSQ(0)	0.087***	0.043***	0.036***	0.021***	0.032***
	(11.09)	(5.59)	(4.02)	(3.73)	(3.46)
RETSQ(−1)	0.021***	0.022***	0.006	0.018**	0.022**
	(3.07)	(3.15)	(1.10)	(2.18)	(2.48)
RETSQ(−2)	0.014**	0.010*	0.027***	0.020**	0.011*
	(2.18)	(1.88)	(3.25)	(2.18)	(1.72)
RETSQ(−3)	0.005	0.025***	0.011	0.012*	−0.001
	(0.95)	(2.94)	(1.48)	(1.91)	(−0.13)
RETSQ(−4)	0.017**	0.016*	0.010	−0.005	0.006
	(2.23)	(1.93)	(1.64)	(−0.65)	(0.86)
VLTY	0.010	0.010	0.006	0.025	0.067**
	(0.38)	(0.29)	(0.18)	(0.70)	(2.15)
SIZE	0.045	0.060	0.075	0.003	−0.046
	(0.59)	(0.62)	(0.79)	(0.03)	(−0.51)
INST	0.496*	0.669**	0.619*	0.671**	0.641**
	(1.80)	(2.26)	(1.90)	(2.15)	(1.97)
COV	0.145*	0.074	0.063	0.003	0.033
	(1.94)	(0.92)	(0.78)	(0.04)	(0.40)
ΔEQVOL	1.215***	0.346***	0.114	0.007	−0.087
	(8.69)	(3.07)	(1.28)	(0.07)	(−0.98)
ΔOPVOL	−0.000	−0.008	−0.008	−0.016	−0.024**
	(−0.01)	(−0.63)	(−0.83)	(−1.59)	(−2.05)
Intercept	−2.250*	−1.891	−2.156	−0.703	0.310
	(−1.69)	(−1.15)	(−1.29)	(−0.45)	(0.20)
R^2 (%)	15.145	13.873	13.572	13.282	13.042

Notes. This table presents summary statistics from daily Fama–MacBeth regressions where the dependent variable is squared realized percent returns. RETSQ(N), MIA, VLTY, SIZE, Log(EQVOL), Log(OPVOL), ΔEQVOL, and ΔOPVOL are defined in the notes of Table 5. RS equals the firm's relative bid-ask spread, OIB equals the absolute order imbalance of buyer- versus seller-initiated volume scaled by total volume, ILLIQ equals the Amihud illiquidity ratio, O/S is the option-to-stock volume ratio, all measured on day t. IV is the option-implied variance, calculated concurrently with MIA as the square of average implied volatility across the firm's closest maturity, at-the-money call and put options, divided by 252 to transform it into a daily measure. SIZE equals the log of market capitalization and VLTY is the standard deviation of returns over the six months ending 10 days prior to day t. INST is the fraction of shares held by institutions and COV is the log of one plus the number of analysts covering the firm. ΔEQVOL and ΔOPVOL equal the percentage change in equity and option volumes relative to the average of each volume over the six months ending 10 days prior to day t. The parentheses contain t-statistics from Fama–MacBeth regressions after Newey–West adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.
***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

IV, other potential proxies for information asymmetry (bid-ask spreads, the Amihud illiquidity ratio, order imbalances, and O/S) are either insignificantly and/or negatively related to future volatility, suggesting that MIA outperforms these prior proxies in capturing the level of information asymmetry among traders associated with the arrival of new information.[14]

4.4. Alternative Implementations and Robustness

In this subsection, we examine alternative implementations of our main tests.

4.4.1. Alternative Estimations of M. As expressed in Equation (20), calculating MIA requires an estimate of $M_{j,t}$, the typical level of O/S for firm j in the absence of private information. In our main tests, we rely on the historical median of O/S as a proxy for $M_{j,t}$. To gauge the sensitivity of our findings to this assumption, we also consider two alternative approaches for calculating $M_{j,t}$. Our first alternative approach defines $M_{j,t}$ as the average level of a firm's O/S in the week after the firm's most recent earnings announcement under the assumption that the public announcement resolves information asymmetry among investors. The use of O/S after earnings announcements yields qualitatively identical results that are untabulated for brevity. Our main tests rely on $M_{j,t}$ computed from the within-firm time series because it is easier to implement empirically and mitigates the impact of cross-sectional differences in the extent to which earnings announcements resolve information asymmetry.

Our second alternative approach defines $M_{j,t}$ based on historically-estimated relations between O/S and firm characteristics using the empirical model from Roll et al. (2010). Using cross-sectional regressions rather than within-firm medians to estimate $M_{j,t}$ refocuses our measure on abnormal O/S relative to firms with similar characteristics rather than firm j's recently observed O/S. This approach mitigates the concern that some firms have persistent informed trade that causes the historical median of their O/S to reflect private information. We implement a version of MIA that uses historically estimated relations between O/S and firm characteristics to estimate $M_{j,t}$, which we refer to as FIA.

To estimate FIA, we use the empirical model from Roll et al. (2010) that estimates the relation between firms' daily O/S and several firm characteristics: log market capitalization, relative spread, implied volatility, option delta, analyst following and forecast dispersion, institutional ownership, and an indicator variable for earnings announcements. We fit daily O/S to these firm characteristics each calendar year and apply the average coefficient estimates from the prior year to firm's current characteristics to estimate $M_{j,t}$.

Panel A Table 7 shows our main findings hold when using FIA in place of MIA. Specifically, FIA is positively associated with relative spreads, the Amihud illiquidity ratio, and order imbalances, and also positively predicts volatility, incremental to the same controls used in Tables 3–6. These findings provide additional support for our central hypothesis that abnormal levels of O/S are indicative of informed trade. Our main tests rely on MIA, rather than FIA, because it is easier to implement empirically and because firms may have persistent differences in their O/S for institutional or behavioral reasons not captured by these chosen characteristics, making their O/S persistently and significantly different from the fitted version of $M_{j,t}$ for reasons other than information asymmetry. The time-series definition of $M_{j,t}$ we use to compute MIA, by contrast, cancels out these persistent differences by using within-firm variation in O/S.

Table 7. Alternative Implementations and Robustness

Dep var:	Eq. Spr.	Op. Spr.	ILLIQ	OIB	RETSQ(1)
	Panel A: Regressions using fitted MIA ("FIA")				
FIA	0.004***	0.001***	0.011***	0.002***	0.719***
	(6.05)	(5.50)	(12.74)	(6.94)	(2.87)
R^2	0.441	0.582	0.642	0.141	0.146
Controls	Yes	Yes	Yes	Yes	Yes
	Panel B: Regressions with firm and year fixed effects				
MIA	0.004***	0.003***	0.019***	0.009***	1.758***
	(8.09)	(24.73)	(11.80)	(38.88)	(9.78)
R^2	0.575	0.622	0.504	0.232	0.070
Controls	Yes	Yes	Yes	Yes	Yes
	Panel C: Regressions excluding options expiration weeks				
MIA	0.011***	0.002***	0.028***	0.016***	1.702***
	(8.59)	(6.77)	(6.48)	(22.96)	(5.70)
R^2	0.464	0.587	0.607	0.139	0.112
Controls	Yes	Yes	Yes	Yes	Yes

Notes. Panel A presents summary statistics from daily Fama–MacBeth regressions where the dependent variables are Eq. Spr., the relative bid-ask spread for the firm's equity; Op. Spr., the relative bid-ask spread for the firm's options, ILLIQ, the Amihud illiquidity ratio; OIB, the absolute order imbalance of buyer- versus seller-initiated volume scaled by total volume; and RETSQ(1), the square of the firm's raw return. Eq. Spr., Op. Spr., ILLIQ, and OIB are measured the same day t as our multimarket information asymmetry measure FIA, whereas RETSQ(1) is measured one trading day later. FIA is an estimate of the fraction of market participants that receive a private signal regarding a firm's future dividends that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. The M used to compute FIA is the fitted value from the cross-sectional regression of O/S on firm characteristics suggested in Roll et al. (2010). Panel B contains panel regression results that include firm and year fixed effects. Panel C presents summary statistics from daily Fama–MacBeth regressions of our main outcome variables on MIA, where we exclude dates coinciding with options expiration weeks. Control variables are used throughout but omitted from the tables for brevity. The parentheses contain t-statistics from Fama–MacBeth regressions after Newey–West adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

4.4.2. Robustness to Alternative Implementations. As noted in Section 3, a potential concern is that MIA captures cross-sectional differences in the volatility of uninformed trade, σ_u, that persist over time, rather than differences in informed trade. To address this concern, and other potential persistent cross-sectional differences, panel B of Table 7 replicates our main tests using panel regressions that include firm and year fixed effects. These tests yield similar results for all three outcome variables, suggesting that MIA also serves to identify within-firm variation in information asymmetry.

Additionally, in panel C of Table 7 we present our main tests when excluding observations that fall in option expiration weeks. We show that our main findings are robust to this specification, which mitigates concerns that our findings are driven by mechanical trading volume associated with option traders rolling forward to the next expiration date. In untabulated tests, we also find that our results are similarly robust to the exclusion of volume during option expiration weeks in the calculation of both O/S and M; however, we do not implement this exclusion in our main tests because our goal is to deliver a measure of informed trade that is both easy to implement and available for the largest possible sample of firm-days.

4.5. MIA as a Conditioning Variable

In this section, we evaluate MIA as a measure of information asymmetry by examining whether it serves as an ex ante conditioning variable that helps distinguish between informed and uninformed sources of price pressure. In our first set of tests, we examine whether the returns to short-term return reversal strategies vary as a function of MIA. Based on the models and evidence in Hendershott and Seasholes (2007), Nagel (2012), and Hendershott and Menkveld (2014) that feature risk-averse market makers, our tests are motivated by the idea that daily returns reverse regardless of the level of information asymmetry but reverse more substantially when MIA is low because these returns are more likely to reflect price pressure from uninformed traders. Conversely, we expect that daily returns reversals are weaker when MIA is high because these returns are more likely to reflect informed traders impounding information into prices.

In Table 8, we examine the returns of a daily reversal strategy after conditioning on MIA. For each trading day, we independently sort stocks into quintile portfolios by market-adjusted returns as well as MIA on day t. To avoid bid-ask bounce and potential look-ahead bias from using summary option volumes, we skip one trading day and measure the returns to each portfolio on day $t+2$. The reversal portfolio consists of a long position in stocks in the lowest quintile of returns on day t and a short position stocks in the highest quintile. Within each MIA quintile, we regress the reversal portfolio's returns on five contemporaneous risk factors: the three Fama and French (1993) factors (MKTRF, SMB, and HML), the momentum factor (UMD), and a liquidity factor (LIQ). The intercept (INT) is the portfolio alpha. The Fama–French and momentum factors are downloaded from Ken French's website,[15] and the liquidity factor is computed as the daily return of a strategy long firms in the highest decile of prior-month average ILLIQ and short firms in the lowest decile.

Panel A of Table 8 shows that portfolio alphas are significantly positive across all MIA portfolios. More importantly, the portfolio alpha for the lowest MIA quintile (15.0 basis points, t-statistic = 6.08) is more than double the alpha for the highest MIA quintile (6.5 basis points, t-statistic = 2.60). The average daily difference is 8.5 basis points and statistically significant (t-statistic = 3.37), indicating that returns exhibit significantly more reversal among stocks with less information asymmetry.

Prior research demonstrates that short-term return reversals are larger when equity turnover is higher (e.g., Campbell et al. 1993, Llorente et al. 2002, Avramov et al. 2006). To mitigate concerns that MIA predicts variation in return reversals through its association with equity turnover, panel B of Table 8 examines daily return reversals in a multivariate context when controlling for the impact of equity turnover. We measure equity turnover, EQTO, on day t as total share volume scaled by total shares outstanding. After assigning firms to quintiles of day t returns, which we refer to as QRET, we run Fama–MacBeth daily regressions where the dependent variable is the firm's market-adjusted return on day $t+2$.

Across columns (1)–(4), the coefficient on QRET is significantly negative (t-statistics ranging from −6.46 to −8.05), indicating that daily returns predictably reverse. We also find that the interaction term between QRET and MIA is significantly positive (t-statistics ranging from 3.47 to 4.23), which is consistent with our findings in panel A that returns exhibit smaller reversals when information asymmetry is high. The interaction effect between QRET and MIA is robust in controlling for the interaction effect between MIA and equity turnover.[16]

Column (4) shows that the interaction effect between MIA and QRET is also robust in controlling for firm size, book-to-market, and momentum, and is distinct from the return predictability associated with the O/S ratio as documented in Johnson and So (2012). Finally, column (5) documents that the interaction effect between MIA and QRET is distinct from the interaction effect between firm size and QRET, which provides evidence that MIA captures variation in information asymmetry incremental to coarser proxies such as firm size.

The results in Table 8 also speak to the debate about whether information asymmetry is related to expected returns. Specifically, column (1) shows an insignificant relation between MIA and future stock returns. The negative coefficient on MIA in columns (2)–(4) is due to the positive interaction effect between QRET and MIA. These results are consistent with the conclusion in Duarte and Young (2009) and Mohanram and Rajgopal (2009) that information asymmetry is not priced in the cross section of stock returns. Even in the absence of a direct relation between MIA and future returns, the results in Table 8 not only support the use of MIA as a proxy for information asymmetry, but also attest to the practical use of MIA as a conditioning variable in short-term reversal strategies.

Table 8. Reversal Strategy Returns Conditioning on MIA

Panel A: Daily return reversal strategy factor loadings

	INT	MKTRF	SMB	HML	UMD	LIQ
Q1: Low MIA	0.150	0.286	0.086	−0.291	0.027	0.163
	(6.08)	(7.59)	(2.02)	−(7.17)	(1.00)	(3.88)
Q2	0.179	0.218	0.104	−0.181	0.036	0.081
	(7.38)	(5.88)	(2.50)	−(4.54)	(1.34)	(1.96)
Q3	0.131	0.244	0.054	−0.231	0.061	0.114
	(5.45)	(6.63)	(1.31)	−(5.85)	(2.28)	(2.78)
Q4	0.133	0.240	0.111	−0.233	0.040	0.131
	(5.56)	(6.55)	(2.68)	−(5.90)	(1.49)	(3.19)
Q5: High MIA	0.065	0.275	0.150	−0.200	0.017	0.158
	(2.60)	(7.20)	(3.50)	−(4.87)	(0.60)	(3.71)
High − Low MIA	0.085	0.011	−0.065	−0.091	0.011	0.005
	(3.37)	(0.29)	−(1.50)	−(2.19)	(0.38)	(0.12)

Panel B: Daily Fama–MacBeth reversal regressions

	(1)	(2)	(3)	(4)	(5)
QRET	−0.035***	−0.046***	−0.038***	−0.044***	−0.020
	(−7.20)	(−7.83)	(−6.46)	(−8.05)	(−1.19)
QRET × MIA	—	0.025***	0.024***	0.023***	0.019***
	—	(4.23)	(4.14)	(4.08)	(3.47)
MIA	0.008	−0.042**	−0.041**	−0.031**	−0.023
	(0.68)	(−2.48)	(−2.46)	(−2.12)	(−1.62)
QRET × EQVOL	—	—	−0.000***	−0.000***	−0.000***
	—	—	(−3.01)	(−2.70)	(−2.67)
EQVOL	—	—	0.000	0.000	0.000
	—	—	(1.57)	(1.58)	(1.63)
MOMEN	—	—	—	0.000	0.000
	—	—	—	(0.73)	(0.75)
SIZE	—	—	—	0.004	0.007
	—	—	—	(0.97)	(1.43)
LBM	—	—	—	−0.020	−0.020
	—	—	—	(−0.85)	(−0.86)
O/S	—	—	—	−0.001***	−0.001***
	—	—	—	(−2.73)	(−2.77)
QRET × SIZE	—	—	—	—	−0.001
	—	—	—	—	(−1.29)
Intercept	0.069***	0.089***	0.073***	0.026	−0.021
	(4.51)	(5.49)	(4.99)	(0.32)	(−0.24)
R^2 (%)	1.587	1.796	3.548	7.078	7.295

Notes. Panel A presents alphas on day $t+2$ for portfolios double-sorted by returns and MIA on day t. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. Within each MIA quintile, we compute the return of a portfolio with long positions in firms in the lowest return quintile and short positions in firms in the highest return quintile. Portfolio returns are measured on day $t+2$ to mitigate look-ahead bias and bid-ask bounce, and regressed on five contemporaneous risk factors: the three Fama–French factors (MKTRF, SMB, and HML), the momentum factor (UMD), and a liquidity factor (LIQ). The intercept in this regression (INT) is the portfolio alpha. All returns are shown as percentages, and t-statistics are shown in parentheses. Panel B presents summary statistics from daily Fama–MacBeth regressions where the dependent variable is the raw return on day $t+2$. QRET is the quintile rank of returns on day t, ranging in value from 1 to 5. EQTO is a firm's total equity volume scaled by total shares outstanding on day t. SIZE equals the log of market capitalization, LBM equals the log of one plus the ratio of book to market equity, and O/S is the option-to-stock volume ratio on day t. MOMEN is the cumulative market-adjusted return over the six months ending 10 days prior to t. The parentheses contain t-statistics from the Fama–MacBeth regressions after Newey–West adjustments for autocorrelation up to 25 lags. The sample consists of 3,533,826 firm-days spanning from 1996 through 2013.

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

To further examine whether MIA distinguishes informed from uninformed sources of price pressure, Table 9 examines the returns to portfolios double-sorted by MIA and implied-volatility spreads. Following Cremers and Weinbaum (2010), we calculate implied volatility spreads as the difference in implied-volatility between at-the-money call and put options of the same strike price and expiration date. Cremers and Weinbaum (2010) show that IV spreads positively predict equity returns, suggesting that positive

Table 9. Implied-Volatility-Spread Strategy Returns Conditioning on MIA

	INT	MKTRF	SMB	HML	UMD	LIQ
			Panel A: Strategy factor loadings (full sample)			
Q1: Low MIA	0.058	0.010	−0.092	−0.141	0.023	0.021
	(3.33)	(0.38)	−(3.10)	−(4.94)	(1.19)	(0.71)
Q2	0.064	0.035	−0.068	−0.079	0.018	0.041
	(3.72)	(1.32)	−(2.29)	−(2.77)	(0.97)	(1.38)
Q3	0.039	0.008	−0.039	−0.151	−0.010	−0.025
	(2.25)	(0.30)	−(1.31)	−(5.35)	−(0.53)	−(0.84)
Q4	0.080	−0.027	0.001	−0.138	−0.007	−0.008
	(4.86)	−(1.09)	(0.03)	−(5.08)	−(0.36)	−(0.28)
Q5: High MIA	0.119	0.044	−0.030	−0.163	−0.015	0.074
	(7.10)	(1.71)	−(1.03)	−(5.90)	−(0.79)	(2.57)
High – Low MIA	0.062	0.034	0.063	−0.022	−0.038	0.053
	(2.59)	(0.94)	(1.53)	−(0.56)	−(1.43)	(1.30)
			Panel B: Strategy factor loadings (O/S > M subsample)			
Q1: Low MIA	0.048	−0.002	−0.101	−0.122	0.035	−0.002
	(2.02)	−(0.05)	−(2.48)	−(3.14)	(1.35)	−(0.06)
Q2	0.078	0.055	−0.016	−0.081	0.017	0.039
	(3.14)	(1.45)	−(0.38)	−(1.97)	(0.62)	(0.91)
Q3	0.035	0.000	−0.050	−0.155	−0.019	−0.002
	(1.39)	(0.01)	−(1.15)	−(3.73)	−(0.68)	−(0.05)
Q4	0.122	0.001	0.018	−0.146	−0.033	0.033
	(5.05)	(0.02)	(0.44)	−(3.65)	−(1.23)	(0.80)
Q5: High MIA	0.150	−0.012	−0.125	−0.178	−0.035	−0.022
	(4.64)	−(0.23)	−(2.24)	−(3.31)	−(0.98)	−(0.40)
High – Low MIA	0.107	−0.013	−0.027	−0.072	−0.076	−0.025
	(2.65)	−(0.22)	−(0.38)	−(1.07)	−(1.71)	−(0.36)
			Panel C: Strategy factor loadings (O/S < M subsample)			
Q1: Low MIA	0.063	0.028	−0.055	−0.146	0.008	0.053
	(2.51)	(0.73)	−(1.28)	−(3.52)	(0.30)	(1.22)
Q2	0.052	−0.007	−0.061	−0.088	0.028	0.025
	(2.15)	−(0.20)	−(1.47)	−(2.21)	(1.05)	(0.62)
Q3	0.036	0.011	0.006	−0.156	−0.008	−0.039
	(1.39)	(0.28)	(0.14)	−(3.65)	−(0.29)	−(0.88)
Q4	0.053	−0.068	−0.024	−0.073	0.024	−0.044
	(2.17)	−(1.83)	−(0.58)	−(1.84)	(0.88)	−(1.07)
Q5: High MIA	0.115	0.024	−0.002	−0.094	−0.002	0.042
	(5.21)	(0.70)	−(0.06)	−(2.59)	−(0.09)	(1.12)
High – Low MIA	0.050	−0.003	0.052	0.053	−0.009	−0.011
	(1.49)	−(0.06)	(0.91)	(0.96)	−(0.24)	−(0.18)

Notes. This table presents alphas on day $t+2$ for portfolios double-sorted by the implied-volatility spread and MIA on day t. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. The implied-volatility spread is the difference in implied volatility across at-the-money call and put options of the same strike price and expiration date. Within each MIA quintile, we compute the return of portfolio long firms in the highest implied-volatility-spread quintile and short firms in the lowest implied-volatility-spread quintile. Portfolio returns are measured on day $t+2$ to mitigate look-ahead bias and bid-ask bounce, and regressed on five contemporaneous risk factors: the three Fama–French factors (MKTRF, SMB, and HML), the momentum factor (UMD), and a liquidity factor (LIQ). The intercept in this regression (INT) is the portfolio alpha. Panel A shows results for the full sample, panel B shows results for observations where the option-to-stock volume ratio (O/S) is greater than the historical median M, and panel C shows the results for observations where the option-to-stock volume ratio (O/S) is less than the historical median M. All returns are shown as percentages, and t-statistics are shown in parentheses.

(negative) implied-volatility spreads reflect price pressure in options markets from informed traders with good (bad) news. Based on this interpretation, we predict that IV spread strategy returns should be stronger following high $\text{MIA}_{j,t}$ days, particularly those with $O/S > M$, because the price pressure is more likely to stem from informed traders.

In Table 9, we assign firms to quintile portfolios formed on IV spreads on day t and again skip one trading day by examining returns to each portfolio on day $t+2$. The IV spread strategy consists of long positions in the highest quintile and short positions in the lowest quintile of IV spread. The portfolio alphas reported in panel A are significantly positive across all MIA portfolios, consistent with the general pattern documented in Cremers and Weinbaum (2010). Additionally, the portfolio alpha for the lowest MIA quintile is 5.8 basis points (t-statistic = 3.33) whereas the alpha for the highest MIA quintile is 11.9 basis points (t-statistic = 7.10). The average daily difference in alphas across high and low MIA portfolios is 6.2 basis points and statistically significant (t-statistic = 2.59), indicating that implied-volatility spreads are more likely to reflect directional price pressure from informed trade when information asymmetry is high.

In panels B and C of Table 9, we partition our sample based on whether the O/S ratio is abnormally high (i.e., O/S > M) or low (i.e., O/S < M) and show that the interaction effect between the IV spread and MIA is concentrated in cases where O/S > M. This evidence is consistent with abnormally high O/S being indicative of a higher concentration of informed trading in options markets.

Together, the results of Tables 8 and 9 further support the use of MIA as a dynamic measure of information asymmetry, while also mitigating the concern that MIA primarily reflects volatility or vega trading.

5. Additional Analyses

5.1. Information Asymmetry Around Information Events

We conduct additional tests that gauge the validity of MIA as a measure of information asymmetry among traders by examining how MIA varies around two firm-specific information events, earnings announcements and Form 8-K filings with the U.S. Securities and Exchange Commission (SEC). We predict that information asymmetry rises prior to information events and subsequently declines as private information is announced and becomes public. To test this prediction, we obtain quarterly earnings announcement dates from Compustat and create a new data set consisting of 63,310 earnings announcements spanning from 1996 through 2013. We also obtain a data set of 648,387 unique firm-days spanning from 1996 through 2011 on which a firm filed an 8-K, downloaded from the SEC's EDGAR website.[17]

The SEC requires firms to disclose any material event within four business days via Form 8-K, meaning the timing of 8-K filings is driven by the arrival of fundamental news to the firm rather than a pre-determined schedule. Further details about Form 8-K filings and data we use can be found in Niessner (2015). For each event, we calculate abnormal MIA on trading days $d-10$ through $d+10$, where d is the event date, by subtracting the firm's average MIA calculated from $d-41$ through $d-11$ from MIA and scaling by the standard deviation of MIA over the same window.

Figure 3 confirms that MIA follows a remarkably similar pattern around firms' earnings announcements and Form 8-K filing dates. The first panel presents the average of abnormal MIA across firm-quarters in each trading day surrounding firms' quarterly earnings announcements. For each calendar quarter, we calculate the cross-sectional average of abnormal MIA for each day relative to the announcement date and then take the time-series average over the sample window. The second panel presents analogous estimates of abnormal MIA in the trading days surrounding firms' 8-K filing dates. In both panels, MIA rises dramatically prior to event dates, which is consistent with information leakage and an increased fraction of market participants trading on private information ahead of the event. MIA then sharply declines on the day of the event, consistent with the information event removing or reducing the informed agents' informational advantage. Moreover, MIA decreases further the day after the event, remains low for multiple days in the post-event period, and gradually rises back to pre-event levels by day $d+10$.

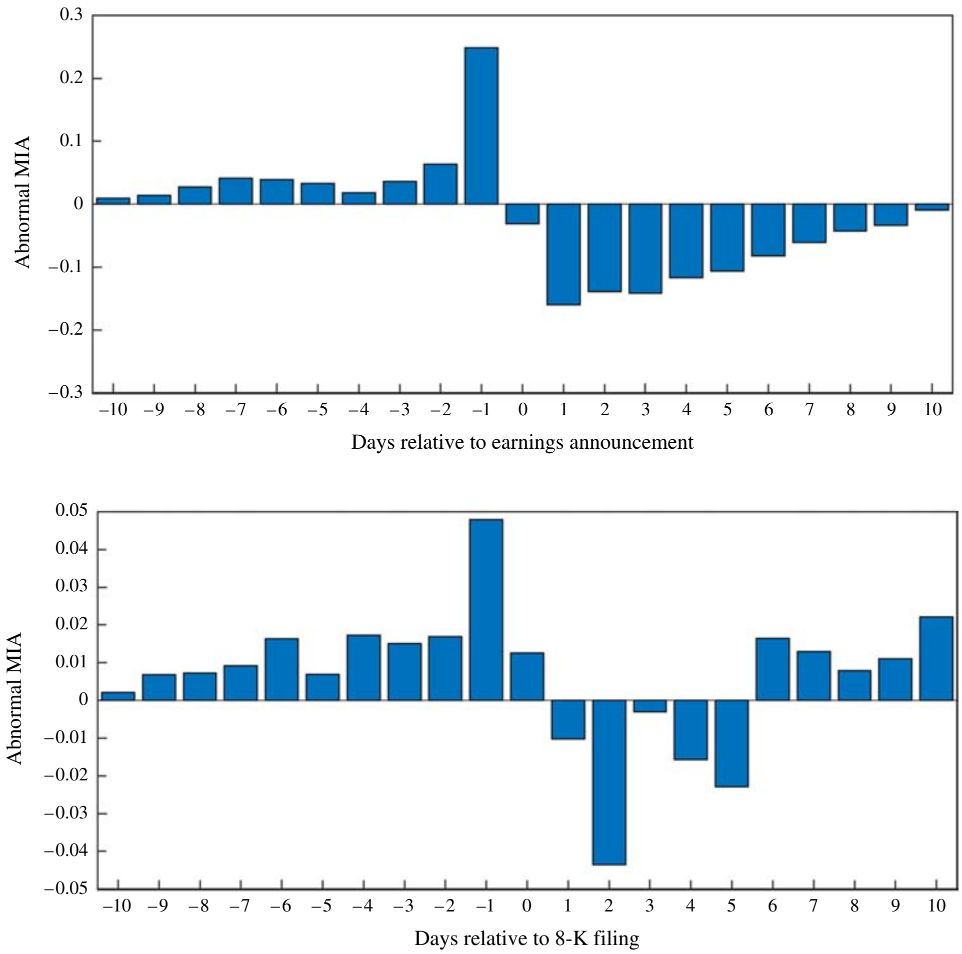
Both the pre-event rise and the post-event drop in MIA are stronger for earnings announcements than 8-K filings, consistent with fewer investors having private information about 8-K filings because they are less likely to be pre-scheduled events. However, to the extent that 8-K events are unscheduled, the observed pattern in MIA is all the more surprising, suggesting that MIA detects informed trade even in cases in which the event is not scheduled or publicly disclosed in advance. Together, the results in Figure 3 highlight a dramatic shift in MIA surrounding information events. In doing so, our findings attest to the benefits of MIA as a daily measure by highlighting the dynamics of information asymmetry in a way that is not feasible with lower frequency proxies such as PIN or analyst coverage.

5.2. Comparing MIA and PIN

Our next analyses examine the relation between PIN and MIA after aggregating daily observations of MIA to the quarterly level by taking the median MIA for all observations in a firm-quarter. We use PIN estimates as calculated in Brown et al. (2009), which in untabulated tests we find have a 41.4% correlation with median MIA.

To compare the effectiveness of these two alternative measures of information asymmetry, Table 10 reports cross-sectional means of quarterly MIA and PIN before and after exogenous terminations of analyst coverage. We predict that MIA should rise along with information asymmetry following an exogenous reduction in analyst coverage. There are two complimentary reasons this should occur. The first, as emphasized in Kelly and Ljungqvist (2012), is that uninformed traders become less active because they rely on analysts to process information. The second

Figure 3. (Color online) Changes in MIA Surrounding Information Events



Notes. This figure presents the time-series average of abnormal MIA in the trading days surrounding firms' quarterly earnings announcements and 8-K filings. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. We calculate abnormal MIA by subtracting the average MIA calculated from $t-41$ to $t-11$ and scaling by the standard deviation of MIA over the same window, where t is the event date. We obtain quarterly earnings announcement dates from Compustat. For each calendar quarter, we calculate the average abnormal MIA for each day relative to the earnings announcement date and then take the time-series average over our sample, which consists of 63,310 earnings announcements spanning from 1996 through 2013. We repeat this exercise using a data set of 648,387 unique firm-days spanning from 1996 through 2011 on which a firm filed an 8-K, which was downloaded from the SEC's EDGAR website and graciously provided to us by Marina Niessner.

possibility is that informed traders' information advantages increases with the reduction in analyst coverage, resulting in more informed trading volume. Both possibilities result in an increase in the fraction of traders with private information, and so an effective proxy for information asymmetry should increase following these exogenous shocks to analyst coverage.

We identify exogenous reductions in analyst coverage driven by closures of brokerage houses as studied in Kelly and Ljungqvist (2012). Following Kelly and Ljungqvist (2012), we refer to firms with exogenous coverage terminations as "treatment" firms. For each treatment firm, we form a control group by selecting stocks with the same size and book-to-market quintile assignment in the quarter prior to termination, subject to the requirement that they were covered by one or more analysts and were not themselves subject to a termination in the quarter before or after the event for the treatment firm. When there are more than five matches, we select the five stocks that are closest to the treatment firm in terms of the pre-event level of the asymmetric information proxy (i.e., MIA or PIN). For each treatment firm j, we construct the percentage difference-in-difference estimate of changes in the asymmetric information proxy as $[(post_j - pre_j)/pre_j] - [(\overline{post}_{control} - \overline{pre}_{control})/\overline{pre}_{control}]$ as well as the level difference-in-difference estimate as $[(post_j - pre_j)] - [(\overline{post}_{control} - \overline{pre}_{control})]$. To identify firms where the loss of analyst coverage is likely to significantly impact information asymmetry, we limit our sample to firms with no more than five analysts prior to the coverage termination, resulting in 556 firm-quarter observations spanning from 2000 through 2008.

Table 10. Coverage Terminations and Information Asymmetry

	Terminations		Matched controls		Percent DiD		Level DiD	
	Before	After	Before	After	Mean	p-value	Mean	p-value
MIA	0.451	0.521	0.461	0.499	0.105	0.011	0.032	0.056
O/S	6.201	5.765	6.744	5.841	0.191	0.104	0.467	0.446
PIN	0.109	0.116	0.113	0.120	0.049	0.222	0.000	0.932

Notes. This table reports cross-sectional means of MIA, O/S, and PIN before and after exogenous terminations of analyst coverage. MIA is an estimate of the fraction of volume originating from informed traders that we calculate for each firm-day using Equation (20) and observed trading volumes in the firm's options and stock. We calculate MIA at the quarterly level by taking the median value over all observations of a given firm in the specified calendar quarter. O/S is the option-to-stock volume ratio. PIN is the probability of informed trade as calculated in Brown et al. (2009). We identify exogenous reductions in analyst coverage driven by acquisitions and closures of brokerage houses as studied in Kelly and Ljungqvist (2012). Following Kelly and Ljungqvist (2012), we refer to firms with exogenous coverage terminations as "treatment" firms. For each treatment firm, we form a control group by selecting stocks with the same size and book-to-market quintile assignment in the quarter prior to termination, subject to the requirement that they were covered by one or more analysts and were not themselves subject to a termination in the quarter before or after the event for the treatment firm. When there are more than five matches, we select the five stocks that are closest to the treatment firm in terms of the pre-event level of the asymmetric information proxy (i.e., MIA or PIN). For each treatment firm j, we construct the percentage difference-in-difference estimate of changes in the asymmetric information proxy as $[(post_j - pre_j)/pre_j] - [(\overline{post}_{control} - \overline{pre}_{control})/\overline{pre}_{control}]$ as well as the level difference-in-difference estimate as $[(post_j - pre_j)] - [(\overline{post}_{control} - \overline{pre}_{control})]$. The sample consists of 556 firm-quarter observations spanning from 2000 through 2008.

The results in Table 10 show that MIA significantly increases following coverage terminations relative to matched control firms. These findings demonstrate that MIA detects exogenous shocks to information asymmetry driven by changes in analyst coverage. Table 10 also shows average O/S does not significantly increase following exogenous coverage terminations, meaning that increases in MIA are not driven by a shift in the level of O/S unrelated to changes in information asymmetry. Subjecting quarterly estimates of PIN to the same tests, we also find no evidence that PIN changes surrounding coverage terminations. These results indicate that although MIA and PIN appear to measure related constructs, only MIA detects exogenous shocks to information asymmetry.

6. Conclusion

The central contribution of this paper is the development of a new proxy for information asymmetry among traders that leverages how trades are dispersed across equity and options markets. The primary innovation of our approach is that we study informed trade in a multimarket setting. The addition of a second market allows us to proxy for the fraction of traders with an information advantage using abnormal volume imbalances across the two markets, as opposed to the approach used in prior research that relies on imbalances between buyer- versus seller-initiated trades. Our multimarket measure of informed trade, MIA, is simple to calculate empirically because it relies only on volume totals, does not require identifying the direction of trades, does not entail estimating a structural model, and can be estimated at the daily level.

We implement MIA empirically and show that it performs remarkably well in response to a battery of empirical tests designed to assess its validity as a measure of information asymmetry among traders. Specifically, we show that MIA is positively associated with bid-ask spreads, price impact, and order imbalances, offers significant predictive power for future volatility, and distinguishes between informed and uninformed sources of price pressure. We also show that MIA rises before firms' earnings announcements and 8-K filing dates, and falls immediately afterward. Finally, MIA detects increases in information asymmetry driven by exogenous reductions in analyst coverage. Taken together, our results indicate that MIA has many desirable properties as a measure of the degree of information asymmetry among traders, as well as many practical benefits as a predictor of volatility, liquidity, and the returns to short-term trading strategies.

Acknowledgments

For helpful comments, the authors thank Lauren Cohen (department editor); the anonymous referees; discussants Kevin Crotty, Terrence Hendershott, and Eric Hughson; Yonca Ertimur, Yesol Huh, Harvey Jing, Charles Lee, Sergey Maslennikov, Stefan Nagel, Jun Pan, Suhas Sridharan, Kumar Venkataram, and Wei Wu; and seminar participants at Massachusetts Institute of Technology, the University of Texas at Austin, the University of Texas at San Antonio, the 2013 Lone Star Conference, the 2013 Financial Research Association (FRA) Conference, and the 2014 Finance Down Under Conference. The authors also thank Marina Niessner for graciously sharing 8-K data and Stephen Brown for sharing quarterly PIN data.

Endnotes

[1] See Section 2 for further discussion of the PIN measure and its relation to our work.

[2] Related empirical evidence in Aktas et al. (2007), Benos and Jochec (2007), and Collin-Dufresne and Fos (2015) indicates PIN is abnormally low when information asymmetry should be at its highest.

[3] In Back (1993) and most other models using the Kyle (1985) modeling framework, there is always an informed trader ($\phi = 1$). We add

the parameter ϕ as a natural source of exogenous variation in the extent of information asymmetry and test whether our MIA measure detects this variation.

[4] To simplify the solution, market makers do not condition prices on order flow in the other assets. Allowing market makers to condition prices on simultaneous demands in all three markets would make it harder for informed traders to reduce price impact by trading multiple assets, thereby causing them to further concentrate their demands in a single market, strengthening our results.

[5] Selling options when v is near zero can be thought of as a form of volatility trading because the informed trader knows big changes in the stock price are unlikely. Volatility trading of this type is inevitable in any model with directional private information because the informed trader's posterior estimate of the stock's value will always have lower variance than an uninformed prior. In this sense, a directional signal is always also negative volatility signal.

[6] We parameterize σ_u to be sufficiently small that $\tilde{u}$ is between 0 and 1 throughout our simulated samples.

[7] A potential concern is that Table 1 shows that average MIA is increasing in σ_u despite no change in average θ, meaning our empirical results could be driven in part by cross-sectional differences in σ_u. We address this possibility by showing our results are robust to controlling for firm fixed effects.

[8] http://wrds.wharton.upenn.edu/.

[9] We omit the prior 10 trading days from our calculation of M to reduce the influence of long-lived private information on our calculation of the option-to-stock volume ratio in the absence of private information. In untabulated tests, our results do not appear sensitive to this choice.

[10] In untabulated tests, we find our results are robust to the spread calculation approach developed in Corwin and Schultz (2012).

[11] As discussed in Section 4.4.2, panel regressions that include firm and year fixed effects yield qualitatively similar results, suggesting that MIA also serves to identify within-firm variation in information asymmetry.

[12] In untabulated tests, we find that MIA also predicts alternative volatility metrics, including the intraday high-low spread from Ni et al. (2008) and intraday realized volatility from Corwin and Schultz (2012).

[13] In untabulated results, we again disaggregate MIA into cases where O/S increases and decreases relative to M and find that only increases predict future volatility incremental to IV. These findings suggest that IV is more likely to reflect the information content of decreases in option volume for future volatility.

[14] ILLIQ, OIB, and O/S are positively related to squared returns in untabulated univariate regressions, but negatively related in the horse race in Table 6, perhaps because they are subsumed by realized and implied volatilities.

[15] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[16] We find qualitatively similar results for return reversals in weekly data. These results are untabulated but available upon request.

[17] https://www.sec.gov/edgar/searchedgar/companysearch.html.

References

Aktas N, De Bodt E, Declerck F, Van Oppens H (2007) The PIN anomaly around M&A announcements. *J. Financial Markets* 10(2):169–191.

Anthony JH (1988) The interrelation of stock and options market trading-volume data. *J. Finance* 43(4):949–964.

Asquith P, Oman R, Safaya C (2010) Short sales and trade classification algorithms. *J. Financial Markets* 13(1):157–173.

Avramov D, Chordia T, Goyal A (2006) Liquidity and autocorrelations in individual stock returns. *J. Finance* 61(5):2365–2394.

Back K (1993) Asymmetric information and options. *Rev. Financial Stud.* 6(3):435–472.

Benos E, Jochec M (2007) Testing the PIN variable. Working paper, University of Illinois at Urbana–Champaign, Champaign.

Black F (1975) Fact and fantasy in the use of options. *Financial Analysts J.* 31(4):36–72.

Boehmer E, Grammig J, Theissen E (2007) Estimating the probability of informed trading—does trade misclassification matter? *J. Financial Markets* 10(1):26–47.

Brown S, Hillegeist SA, Lo K (2009) The effect of earnings surprises on information asymmetry. *J. Accounting Econom.* 47(3): 208–225.

Campbell JY, Grossman SJ, Wang J (1993) Trading volume and serial correlation in stock returns. *Quart. J. Econom.* 108(4):905–939.

Cao C, Chen Z, Griffin JM (2005) Informational content of option volume prior to takeovers. *J. Bus.* 78(3):1073–1109.

Chan K, Chung YP, Fong W-M (2002) The informational role of stock and option volume. *Rev. Financial Stud.* 15(4):1049–1075.

Collin-Dufresne P, Fos V (2015) Do prices reveal the presence of informed trading? *J. Finance* 70(4):1555–1582.

Corwin SA, Schultz P (2012) A simple way to estimate bid-ask spreads from daily high and low prices. *J. Finance* 67(2): 719–760.

Cremers M, Weinbaum D (2010) Deviations from put-call parity and stock return predictability. *J. Financial Quant. Anal.* 45(2): 335–367.

Duarte J, Young L (2009) Why is *PIN* priced? *J. Financial Econom.* 91(2):119–138.

Duarte J, Hu E, Young LA (2015) What does the PIN model identify as private information? Working paper, Rice University, Houston.

Easley D, O'Hara M (1987) Price, trade size, and information in securities markets. *J. Financial Econom.* 19(1):69–90.

Easley D, Kiefer NM, O'Hara M (1997) One day in the life of a very common stock. *Rev. Financial Stud.* 10(3):805–835.

Easley D, Lopez de Prado MM, O'Hara M (2012) Flow toxicity and liquidity in a high-frequency world. *Rev. Financial Stud.* 25(5):1457–1493.

Easley D, O'Hara M, Srinivas PS (1998) Option volume and stock prices: Evidence on where informed traders trade. *J. Finance* 53(2):431–465.

Easley D, Kiefer NM, O'Hara M, Paperman JB (1996) Liquidity, information, and infrequently traded stocks. *J. Finance* 51(4): 1405–1436.

Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. *J. Financial Econom.* 33(1):3–56.

Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: Empirical tests. *J. Political Econom.* 81(3):607–636.

Glosten PR, Milgrom LR (1985) Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *J. Financial Econom.* 14(1):71–100.

Grossman SJ, Stiglitz JE (1980) On the impossibility of informationally efficient markets. *Amer. Econom. Rev.* 70(3):393–408.

Hendershott T, Menkveld AJ (2014) Price pressures. *J. Financial Econom.* 114(3):405–423.

Hendershott T, Seasholes MS (2007) Market maker inventories and stock prices. *Amer. Econom. Rev.* 97(2):210–214.

Johnson TL, So EC (2012) The option to stock volume ratio and future returns. *J. Financial Econom.* 106(2):262–286.

Kelly B, Ljungqvist A (2012) Testing asymmetric-information asset pricing models. *Rev. Financial Stud.* 25(5):1366–1413.

Kyle AS (1985) Continuous auctions and insider trading. *Econometrica: J. Econometric Soc.* 53(6):1315–1335.

Lee C, Ready MJ (1991) Inferring trade direction from intraday data. *J. Finance* 46(2):733–746.

Lee C, So EC (2017) Uncovering expected returns: Information in analyst coverage proxies. *J. Financial Econom.* 124(2):331–348.

Llorente G, Michaely R, Saar G, Wang J (2002) Dynamic volume-return relation of individual stocks. *Rev. Financial Stud.* 15(4): 1005–1047.

Mohanram P, Rajgopal S (2009) Is PIN priced risk? *J. Accounting Econom.* 47(3):226–243.

Nagel S (2012) Evaporating liquidity. *Rev. Financial Stud.* 25(7): 2005–2039.

Newey WK, West KD (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3):703–708.

Ni SX, Pan J, Poteshman AM (2008) Volatility information trading in the option market. *J. Finance* 63(3):1059–1091.

Niessner M (2015) Strategic disclosure timing and insider trading. Working paper, Yale University, New Haven, CT.

Odders-White ER, Ready MJ (2008) The probability and magnitude of information events. *J. Financial Econom.* 87(1):227–248.

Roll R, Schwartz E, Subrahmanyam A (2010) O/S: The relative trading activity in options and stock. *J. Financial Econom.* 96(1):1–17.

Stephan JA, Whaley RE (1990) Intraday price change and trading volume relations in the stock and stock option markets. *J. Finance* 45(1):191–220.