The Non-Information Cost of Trading and Its Relative Importance in Asset Pricing†

Kee H. Chung[a,*] and Sahn-Wook Huh[b]

[a] School of Management, University (SUNY) at Buffalo, Buffalo, NY 14260
E-mail: keechung@buffalo.edu

[b] School of Management, University (SUNY) at Buffalo, Buffalo, NY 14260
E-mail: swhuh@buffalo.edu

* *Corresponding Author*

† This paper benefited greatly from the comments and suggestions of Maureen O'Hara (the editor) and two anonymous referees. The authors thank Hank Bessembinder, Andrew Karolyi, Ken Singleton, Avanidhar Subrahmanyam, John Ritter, Petko Kalev, session participants at the 2015 FMA European Conference and the 22nd Annual Conference of the Multinational Finance Society, and colleagues at SUNY-Buffalo for valuable comments and suggestions. The authors also thank Stephen Brown, Soeren Hvidkjaer, Yuxing Yan, Mao Ye, and Lance Young for useful suggestions on PIN estimation; Bidisha Chakrabarty and Andriy Shkilko for valuable discussions about the accuracy of different methods for trade classification; and Seung-Oh Han, Yuxiang Jiang, and Yue Wang for excellent research assistance. We are grateful to Cynthia Cornelius, Tony Kew, and Shawn Mattot at the Center for Computational Research (CCR) for their assistance in addressing various computational issues. Huh gratefully acknowledges the generous financial support from SUNY-Buffalo, the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Ministry of Research and Innovation of Ontario. This work is also supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A3A2036037). All errors are solely the authors' responsibility.

The Non-Information Cost of Trading and Its Relative Importance in Asset Pricing

Abstract

Using intraday order-flow data for a broad and long sample of NYSE/AMEX stocks, we show that the non-information component of trading costs is priced in the cross-section of stock returns. More importantly, we show that the non-information component is much larger and more strongly related to stock returns than the adverse-selection component, indicating that the non-information component plays a more important role in asset pricing than the adverse-section component. We conduct a variety of robustness tests and show that our main results hold for different estimation methods, measures of the adverse-selection cost, study samples, and control variables. We offer plausible explanations for these results.

JEL Classification: G12

Keywords: Order Flows, Lee-Ready Method, Holden-Jacobsen Algorithm, Trading Costs, Order-Processing Cost, Inventory-Holding Cost, Non-Information Cost, Information Asymmetry, Adverse-Section Cost, Asset Pricing, Equilibrium Returns

I. Introduction

Prior research suggests that the cost of trading in a competitive market consists of adverse-selection and non-information components.[1] The adverse-selection (or information-asymmetry) component is the portion of trading costs faced by liquidity providers who trade with informed traders.[2] In contrast, the non-information component is the portion that arises from non-informational reasons such as inventory-holding risks.[3] While numerous studies analyze the effect of the adverse-selection cost on stock returns, relatively little is known about the effect of the non-information component. In this study we analyze the role of the non-information cost of trading and its relative importance in asset pricing using intraday order-flow data for more than 1,800 NYSE/AMEX stocks over the 28-year period from January 1983 to December 2010.

Existing studies seem to agree that illiquidity is priced in the cross-section of stock returns. However, why illiquidity matters in asset pricing is open to debate. Some scholars attribute the cross-sectional pricing of illiquidity to information asymmetry. For instance, Easley and O'Hara (2004) provide a model that underscores the role of asymmetric information in asset pricing and a number of studies provide evidence that is consistent with the prediction of the model (see Section II for a review of these studies). The present study examines whether the non-information cost of trading (e.g., the order-processing and inventory-holding cost) is priced in the cross-section of stock returns, and if so, whether it commands a return premium after controlling for the effect of the adverse-selection component.[4]

Brennan and Subrahmanyam (1996) suggest that the adverse-selection cost plays a more important role in asset pricing than the non-information cost using a sample of two years (1984 and 1988) from the Institute for the Study of Security Markets (ISSM) database. Given the

[1] See Glosten and Harris (1988), Huang and Stoll (1997), Stoll (2000), and references therein.

[2] Liquidity providers impose the adverse-selection component of trading costs on traders to recoup their losses to better-informed traders by selling (buying) a security at a price that is higher (lower) than its value.

[3] Liquidity providers impose the non-information component of trading costs on traders to cover the costs of doing business, such as the inventory-holding and order-processing costs.

[4] Levi and Zhang (2014) investigate the role of these components in event-specific settings such as days before earnings announcements.

considerable changes in market regulation (e.g., Regulation NMS), market structure (e.g., decimalization and market segmentation), trading technologies (e.g., algorithmic trading), and trading behaviors (e.g., order-splitting practice and high-frequency trading) in recent years, our study sheds further light on the relative importance of the non-information and adverse-selection costs using a broader and longer dataset. In addition, prior studies show that the Lee-Ready (1991) method is highly vulnerable to trade classification errors when applied to trade and quote data in high-frequency-trading years. To address this issue, we employ the Holden-Jacobsen (2014) algorithm to match trades and quotes in the last four years (2007-2010) of our sample.

We first show that the non-information cost is priced in the cross-section of stock returns after controlling for the five known pricing factors and other stock attributes. When the non-information component of trading costs is further decomposed into the transitory fixed and variable components, however, only the transitory fixed component is priced. We conduct a variety of robustness tests and show that our main results hold for different estimation models, study samples, control variables, and regression methods.

To assess the magnitudes of the non-information and adverse-selection components of trading costs and their relative importance in asset pricing, we next calculate the proportions of the two components. We find that the non-information component is much larger than the adverse-selection component, confirming the small sample result of Huang and Stoll (1997). We also confirm the finding of earlier studies that the adverse-selection component commands a positive return premium. More importantly and contrary to the small sample result reported in Brennan and Subrahmanyam (1996), we find that stock returns are more strongly related to the non-information component than the adverse-selection component when we include both components in the regressions, indicating that the non-information component plays a more important role in asset pricing than the adverse-selection component.

The non-information cost of trading may play a more important role in asset pricing simply because its absolute magnitude is larger. In addition, it may be easier for investors to measure the non-information component than to measure the adverse-selection component because the former is more closely related to observable stock attributes such as trading volume. The non-information component may be easier to estimate also because it depends on the variables, such as minimum tick sizes, that vary across stocks but do not vary across trades in a given stock. As a result, investors may be more readily incorporate the non-information cost into required returns. We corroborate this explanation by showing that the estimates of the non-information component are less noisy (i.e., smaller standard errors and larger *t*-values) than the estimates of the adverse-selection component.

Several recent studies argue that the role of information asymmetry in asset pricing may not be important to the extent that a large portion of the adverse-selection risk arises from firm-specific idiosyncratic shocks, which is diversifiable by investors and market makers.[5] We find that the positive relation between the probability of informed trading on bad news (*PIN_B*) and stock returns reported in Brennan, Huh, and Subrahmanyam (2015a) remains significant when the non-information cost is included in the regression. In addition, the relation between stock returns and the non-information cost is stronger than the relation between stock returns and *PIN_B*. We interpret this result as evidence that the non-information risk (e.g., risk associated with inventory holding) is less diversifiable than the adverse-selection risk.

The rest of the paper is organized as follows. Section II surveys prior research on the role of the adverse-selection and non-information costs in asset markets. Section III introduces three spread component models, describes data sources, and presents descriptive statistics. Section IV analyzes the effect of the non-information cost on stock returns. In Section V, we conduct additional tests to assess the robustness of our results. Section VI analyzes the magnitude of the adverse-selection and non-information costs of trading and their relative importance in asset pricing. Section VII conducts additional tests using the probability of informed trading (*PIN_B*) and the Amihud illiquidity measure. Section VIII provides a brief summary of the paper.

---

[5] See Hughes, Liu, and Liu (2007), Armstrong, Core, Taylor, and Verrecchia (2011), Lambert, Leuz, and Verrecchia (2012), and Lambert and Verrecchia (2014).

II. Prior Research

A number of studies in the finance and accounting literature examine the role of the adverse-selection risk in financial markets.[6] Easley, Hvidkjaer, and O'Hara (2002) show that stocks with a higher probability of informed trading (*PIN*) provide investors with higher returns. Garleanu and Pedersen (2004) suggest that the bid-ask spread does not directly influence the required return and only the allocation costs associated with adverse selection affect the required return. Chordia, Huh, and Subrahmanyam (2009) show that stocks with a larger Kyle's (1985) lambda (λ) provide higher returns. Kelly and Ljungqvist (2012) show that liquidity is the primary channel through which information asymmetry affects prices. Hwang et al. (2013) show that the adjusted *PIN* purged of a liquidity component (*AdjPIN*) is positively related to the implied cost of equity using data from the Korean stock market. Brennan et al. (2015a, 2015b) analyze the pricing implication of the two components of *PIN* (i.e., good- and bad-news components).

Other studies employ the price impact of a trade (Huang and Stoll, 1997) as a measure of information asymmetry. For example, Bhattacharya et al. (2012) examine the relation between earnings quality and the cost of capital and find evidence that information asymmetry indirectly affects the cost of equity capital. Bhattacharya, Desai, and Venkataraman (2013) test the association between earnings quality and information asymmetry. Levi and Zhang (2014) find that temporary increases in information asymmetry before earnings announcements lead to a higher cost of equity capital.

Duarte and Young (2009) decompose *PIN* into the probability of 'pure' informed trading (*AdjPIN*) and the probability of symmetric order-flow shock (*PSOS*), and show that *PSOS* is positively related to stock returns while *AdjPIN* is not. They also find a high positive correlation between *PSOS* and the Amihud (2002) measure and interpret the result as evidence that *PSOS* captures illiquidity caused by the non-information cost of trading. According to Duarte and

[6] Other studies analyze the pricing implication of liquidity risks. See, for example, Pastor and Stambaugh (2003) and Archarya and Pedersen (2005).

Young (2009), therefore, *PIN* does not capture information asymmetry and *PSOS* solely drives the relation between *PIN* and the cross-section of stock returns. However, Brennan et al. (2015a) find that *PIN* is highly correlated (26%-27%) with the adverse-selection component of the spread measured by the Glosten-Harris (1988) and Foster-Viswanathan (1993) models, but it is weakly correlated (lower than 5%) with the non-information component.[7]

The Amihud (2002) illiquidity measure is defined as the ratio of the absolute return to the dollar trading volume. Thus, it captures the absolute percentage price change per dollar of trading volume, which is, in spirit, analogous to the Kyle's price-impact measure (λ). Indeed, Amihud (2002) shows that his measure is significantly and positively related to Kyle's λ.[8] To the extent that the Amihud illiquidity measure captures the adverse-selection cost, the positive relation between stock returns and *PSOS* documented in Duarte and Young (2009) may reflect, at least in part, the positive cross-sectional relation between stock returns and the adverse-selection cost. In this respect, Duarte and Young (2009) provide limited evidence regarding the relative importance of the adverse-selection and the non-information costs in asset pricing. Interestingly, Lai, Ng, and Zhang (2014) show that stock returns are positively and significantly related to the Amihud (2002) measure, whereas stock returns are *negatively* and significantly related to *PSOS* even when the Amihud measure is not included in the regression. These results raise a question on whether the *PSOS* measure of Duarte and Young (2009) adequately captures illiquidity related to the non-information component, as pointed out in Brennan et al. (2015a).

Prior studies also analyze the role of the non-informational factors (e.g., funding constraints and inventory risk) in liquidity provision and asset pricing. Campbell, Grossman, and Wang (1993) show that returns accompanied by larger trading volume tend to be reversed more strongly and explain this result using a model in which risk-averse market makers are rewarded

---

[7] See Table 2 in Brennan et al. (2015a).

[8] Brennan et al. (2015a) report that the adverse-selection component is positively related to both the original Amihud (2002) measure and its turnover-version [used in Brennen et al. (2013) and Lou and Shu (2014)], with the correlation coefficient ranging from 0.21 to 0.24. In contrast, the non-information component is negatively related to these measures (-0.06 to -0.10), indicating that, contrary to Duarte and Young (2006), the Amihud measure is more likely to capture the adverse-selection cost than the non-information cost.

for accommodating buying or selling pressure from liquidity (non-informational) traders. Pastor and Stambaugh (2003) use the same logic to measure liquidity (i.e., lower liquidity corresponds to stronger volume-related return reversals) and show that stocks that are more sensitive to aggregate liquidity have higher expected returns.

Brunnermeier and Pedersen (2009) provide a model that relates the market liquidity of an asset to traders' funding liquidity and *vice versa*. The authors show that under certain conditions market liquidity and funding liquidity can reinforce each other and predict that speculators' capital is an important determinant of market liquidity and risk premiums. Hendershott and Seasholes (2007) find a negative correlation between specialist inventories and contemporaneous returns and interpret the result as evidence that NYSE specialists are compensated for holding suboptimal portfolios through favorable subsequent price movements. Comerton-Forde et al. (2010) show that temporal variation in asset liquidity depends, at least in part, on the market-maker's financial constraints.

Liu and Wang (2012) develop a theoretical model that emphasizes inventory risks. Nagel (2012) and Hendershott and Menkveld (2014) analyze the pricing effects of excess inventory and limited-risk bearing capacity. Nagel (2012) shows that withdrawal of liquidity supply and associated increase in the expected returns from liquidity provision are main drivers behind the evaporation of liquidity during times of financial market turmoil. Hendershott and Menkveld (2014) find economically significant deviations of prices from fundamental values due to inventory risks born by intermediaries on the NYSE. On the whole, the results of these studies suggest that the non-information cost of trading could potentially play an important role in the equilibrium asset pricing. However, none of these studies provide direct evidence regarding the effect of the non-information cost on the equilibrium stock returns. We provide such evidence in the present study.

III. Estimation of the Non-Information Component of Trading Costs

A. Models of Spread Components

A.1. The Glosten and Harris (1988) Model

As in Kyle (1985) and Admati and Pfleiderer (1988), we assume that the fundamental value (μ_t) of a security evolves as follows:

$$\mu_t = \mu_{t-1} + \lambda S_t V_t + \xi_t; \tag{1}$$

where S_t is equal to $+1$ for buyer-initiated trades and -1 for seller-initiated trades, V_t denotes share or dollar volume at time *t*, and ξ_t represents the public information signal at time *t*.[9]

Glosten and Harris (1988) decompose the total trading cost into the following four components: the transitory fixed cost ($\bar{\varphi}$), the transitory variable cost ($\bar{\lambda}$), the permanent fixed cost (φ), and the permanent variable cost (λ).[10] The first and second components reflect the market-maker rent, inventory-holding cost, and order-processing cost, while the third and fourth components reflect the adverse-selection cost. Glosten and Harris (1988) show that the permanent fixed cost and the transitory variable cost are negligible in their study sample: i.e., $\varphi = \bar{\lambda} = 0$. Hence, we initially assume that $\varphi = \bar{\lambda} = 0$ in our analysis. Given the sign (S_t) of each trade, we can express the observed price, P_t, as follows:

$$P_t = \mu_t + \bar{\varphi} S_t. \tag{2}$$

Plugging Eq. (1) into Eq. (2), we have

$$P_t = \mu_{t-1} + \lambda S_t V_t + \bar{\varphi} S_t + \xi_t. \tag{3}$$

From Eq. (2), we also know that

$$P_{t-1} = \mu_{t-1} + \bar{\varphi} S_{t-1}. \tag{4}$$

Subtracting Eq. (4) from Eq. (3), we can express the price change, ΔP_t, as follows:

$$\Delta P_t = \bar{\varphi}^{GH}(S_t - S_{t-1}) + \lambda^{GH} S_t V_t + \xi_t; \tag{5}$$

[9] Note that μ_t is the expected value of a security, conditional on the information set at time *t*, of a market maker who observes only the order flow, $S_t V_t$, and a public information signal, ξ_t. Thus, the fundamental value of a security at time *t* is determined by the expected value of the security at time *t* - 1 (conditional on the information set at time *t* - 1), signed volume (order flows), and public information.

[10] The transitory component in trading costs reflects the market making costs (e.g., inventory-holding and order-processing costs), while the permanent component reflects the adverse-selection cost. We use the terms the transitory cost and the non-information cost interchangeably throughout the paper. Likewise, we use the information-asymmetry (or permanent) cost and the adverse-selection cost interchangeably.

where $\bar{\varphi}^{GH}$ is the non-information or transitory component of trading costs (i.e., order-processing and inventory-holding costs) and λ^{GH} is the permanent component of trading costs (i.e., adverse-selection cost). We allow for a constant term in Eq. (5) and estimate the cost components for each stock using intraday order flows in each month from January 1983 through December 2010.

A.2. The Foster and Viswanathan (1993) Model

Foster and Viswanathan (1993) use *unexpected* order flows to measure the price impact of trades. Their model accounts for the fact that, if order flows are auto-correlated, only the unpredictable portion of order flows should affect quotes and prices. This approach is compelling, especially because the order-splitting practice in recent years may have caused order flows to be serially correlated. Following Sadka (2006), therefore, we filter order flows by the AR (5) process:

$$S_t V_t = \delta + \sum_{q=1}^{5} k_q S_{t-q} V_{t-q} + \tau_t; \tag{6}$$

where τ_t is the residual from the time-series regression. We measure the unexpected order flows by τ_t and replace $S_t V_t$ in Eq. (5) with τ_t to get the following model:

$$\Delta P_t = \bar{\varphi}^{FV}(S_t - S_{t-1}) + \lambda^{FV}\tau_t + \xi_t'; \tag{7}$$

where $\bar{\varphi}^{FV}$ is the non-information component of trading costs and λ^{FV} is the adverse-selection component.

A.3. The Sadka (2006) Model

A potential drawback of the above two models is that they ignore the permanent fixed cost (φ) and the transitory variable cost ($\bar{\lambda}$). Given that we use a long time-series and broad cross-section of stocks, decomposing trading costs into the four components may provide additional information on the role of different components in asset pricing. As in Sadka (2006), we estimate the unexpected order flows (τ_t), their variance (σ_τ^2), and the fitted value of order flows ($\widehat{S_t V_t}$) from Eq. (6). Next, we obtain the unexpected sign (π_t) of a trade using the following equation:

$$\pi_t = S_t - E_{t-1}(S_t) = S_t - \left\{1 - 2\Phi\left(-\frac{\widehat{S_t V_t}}{\sigma_\tau}\right)\right\};$$

where $\Phi(.)$ denotes the cumulative normal distribution function. We then estimate the four components of trading costs using the following model:

$$\Delta P_t = \bar{\varphi}^S(S_t - S_{t-1}) + \bar{\lambda}^S(S_t V_t - S_{t-1}V_{t-1}) + \varphi^S \pi_t + \lambda^S \tau_t + \varsigma_t; \quad (8)$$

where $\bar{\varphi}^S$ is the transitory fixed cost, $\bar{\lambda}^S$ is the transitory variable cost, φ^S is the permanent fixed cost, and λ^S is the permanent variable cost. We allow for a constant term in Eq. (8) and estimate the four components for each stock using intraday order flow data in each month from January 1983 through December 2010. Note that when $\bar{\lambda}^S = \varphi^S = 0,$ Eq. (8) is reduced to the Foster and Viswanathan (1993) model, specified in Eq. (7) above.

We initially focus on the role of the non-information component in asset pricing. Therefore, we use the parameters that represent the transitory cost (i.e., $\bar{\varphi}^{GH}, \bar{\varphi}^{FV}, \bar{\varphi}^S$, and $\bar{\lambda}^S$) as our key variables. Later in the paper, we also look at how the adverse-selection component measured by λ^{GH} and λ^{FV} is related to stock returns. The results are similar whether we use dollar-volume- or share-volume-based order flows. Hence, we report only the results based on dollar-volume-based order flows.

B. Data Sources, Trade/Quote Matching Methods, and Summary of Order Flow Data

We obtain intraday transaction data from the Institute for the Study of Security Markets (ISSM) for the 1983-1992 period and the NYSE Trades and Automated Quotations (TAQ) for the 1993-2010 period. We limit our study sample to stocks listed on the NYSE and the AMEX because of different trading protocols (Atkins and Dyl, 1997) and data availability for NASDAQ stocks. We use only those stocks with at least 50 trades per month to ensure that the cost estimates from the monthly time-series regressions are reliable. We exclude from the study sample all the trades that are out of sequence, recorded before the open or after the close, or involved in errors or corrections. We also exclude quotes before the open or after the close.

For the sample period from January 1983 to December 2006, we use the Lee and Ready (1991) method to match trades with quotes and classify each trade into a buyer- or seller-initiated trade. We apply the five-second delay rule to match trades with quotes from January 1983 to December 1998. Considering the shorter reporting lag between trades and quotes in later years, we use the two-second delay rule for the 1999-2006 period. Some issues related to the Lee-Ready method and the 'monthly' TAQ (as opposed to the 'daily' TAQ) have been raised by microstructure researchers.[11] However, given that we lag quotes when matching them with trades and use dollar-volume-based (instead of trade-number-based) order flows, the impact of misclassifications on our results is likely to be small for the 1983-2006 period during which high-frequency-trading volume is relatively low.

The past decade has witnessed significant changes in market regulation, market structure, trading technologies, and trading behaviors of market participants. For instance, Stoll (2014) shows that since the mid-2000s the daily number of trades has increased substantially, while the average trade size has decreased, suggesting the prevalence of high-frequency trading (HFT) in recent years, especially from 2007 (see Figure 1 in his study). Arnuk and Saluzzi (2012) argue that Regulation NMS made the speed of execution paramount in the U.S. stock market, thereby triggering a surge of HFT. Given the large HFT volume in recent years, Easley, Lopez de Prado, and O'Hara (2012) and Holden and Jacobsen (2014) suggest that applying the Lee-Ready (1991) method to the monthly TAQ database, which is time-stamped only to the *second* (as opposed to *millisecond*), could induce substantial trade classification errors. To reduce the problem, Holden and Jacobsen (2014) propose a low-cost alternative algorithm,[12] which is applicable to the

[11] Lee and Radhakrishna (2000) and Odders-White (2000) show that the Lee and Ready (1991) method is 85% accurate when applied to the data in the 1990s. O'Hara, Yao, and Ye (2011) argue that order imbalance measures constructed from the TAQ database are subject to errors because the TAQ database excludes odd-lot trades and the errors are particularly severe for measures based on the number of trades. Chakrabarty, Moulton, and Shkilko (2012) report that while the Lee-Ready method's misclassification rates are near zero at the daily aggregate level, they are as high as 30% at the transaction level when contemporaneous quotes are matched with trades, and about 21% when they use the one-second delay rule.

[12] The Holden-Jacobsen (2014) algorithm has the following features: (i) adjustments for withdrawn quotes, (ii) time-interpolation during each one-second period, (iii) matching trades with national best bid and offer (NBBO) quotes across different exchanges, and (iv) excluding crossed or locked NBBOs. So the time delay between quotes and trades is virtually zero (more precisely, milliseconds at most).

monthly TAQ database. The authors show that their algorithm provides more accurate buy/sell classifications (90.4%) than the Lee-Ready (1991) method. Therefore, we use the Holden-Jacobsen (2014) method for the last four years in our sample (2007-2010). After matching trades and quotes based on either of the two methods, if a trade occurs above (below) the prevailing quote midpoint, it is considered buyer-initiated (seller-initiated). To minimize errors, we discard trades executed at the quote midpoints (Sadka, 2006), which constitute 5.79% of the trades when computed with 2007 data.[13]

Table 1 summarizes key attributes of the transaction-level order-flow data used in our study. The total number of trades is more than 17 billion (the same number of bid and ask quotes are matched with trades) over the 28-year study period. By construction, the minimum number of monthly trades (and matched quotes) is 50 for each stock. On average, our sample stocks have 26,980 trades in a month (excluding the trades executed at the quote midpoint). The total number of stock-month observations is 632,614.

C. Descriptive Statistics for the Non-Information Cost Measures

Panel A in Table 2 shows the descriptive statistics for the four non-information cost measures: $\bar{\varphi}_0^{GH}$ ≡ the Glosten and Harris (1988) non-information component; $\bar{\varphi}_0^{FV}$ ≡ the Foster and Viswanathan (1993) non-information component; $\bar{\varphi}_0^{S}$ ≡ the Sadka (2006) transitory fixed cost; and $\bar{\lambda}_0^{S}$ ≡ the Sadka (2006) transitory variable cost.[14] We first calculate the cross-sectional mean, median, standard deviation (*STD*), coefficient of variation (*CV*), skewness, and kurtosis for each variable in each month and then obtain the time-series average of these values over the

[13] To examine whether discarding the trades executed at the quote midpoint causes any significant difference in our results, we estimate the measures in *two* ways using 1,775 component firms that have all 12 monthly estimates in 2007: one estimated with order flows excluding the midpoint trades and the other estimated with order flows including the midpoint trades. The results show that the mean time-series correlation (and cross-sectional correlation) between the two estimates of $\bar{\varphi}_0$'s (excluding vs. including) ranges from 0.998 to 0.999. Similarly, the mean time-series correlation (and cross-sectional correlation) between the two estimates of λ_0's ranges from 0.978 to 0.994.

[14] Following the practice in the literature, we multiply $\bar{\varphi}_0^i$ (and φ_0^i) by 10^2 and $\bar{\lambda}_0^i$ (and λ_0^i) by 10^6 (i = *GH, FV,* or *S*) because these estimates are small.

study period. We calculate the same statistics for their price-scaled values (i.e., $\frac{\bar{\varphi}_0^{GH}}{P}$, $\frac{\bar{\varphi}_0^{FV}}{P}$, $\frac{\bar{\varphi}_0^{S}}{P}$, and $\frac{\bar{\lambda}_0^{S}}{P}$, where P is the previous month-end stock price) and report the results in Panel B.

The two panels show that the distribution of the non-information component estimates is highly leptokurtic and skewed. The large kurtosis suggests that these estimates have many extreme observations. Prior research uses the square-root or logarithmic transformation of the variable to alleviate the influence of extreme observations.[15] However, neither of these transformations is feasible for our component estimates because some of them are negative. We thus Winsorize the raw and price-scaled components at the 0.5th and 99.5th percentiles.

Panel C in Table 2 shows descriptive statistics for the Winsorized values ($\bar{\varphi}^{GH}$, $\bar{\varphi}^{FV}$, $\bar{\varphi}^{S}$, and $\bar{\lambda}^{S}$) of $\bar{\varphi}_0^{GH}$, $\bar{\varphi}_0^{FV}$, $\bar{\varphi}_0^{S}$, and $\bar{\lambda}_0^{S}$ and Panel D shows descriptive statistics for the Winsorized values ($\frac{\bar{\varphi}^{GH}}{P}$, $\frac{\bar{\varphi}^{FV}}{P}$, $\frac{\bar{\varphi}^{S}}{P}$, and $\frac{\bar{\lambda}^{S}}{P}$) of $\frac{\bar{\varphi}_0^{GH}}{P}$, $\frac{\bar{\varphi}_0^{FV}}{P}$, $\frac{\bar{\varphi}_0^{S}}{P}$, and $\frac{\bar{\lambda}_0^{S}}{P}$. We find that the skewness and kurtosis for the Winsorized variables are much smaller, compared to the corresponding non-Winsorized values reported in Panel A and Panel B. Given these results and also to make our study comparable to Sadka (2006) and other previous studies, we use the price-scaled and Winsorized variables as the primary input for our empirical analyses.[16]

To shed some light on the time-series behavior of the non-information component of trading costs, we plot in Figure 1 the monthly value-weighted average values of $\bar{\varphi}^{GH}$, $\bar{\varphi}^{FV}$, and $\bar{\varphi}^{S}$. The figure shows that the non-information cost has declined steadily over time. In particular, we find large drops in the non-information cost at the time of the tick-size reduction in June 1997 (from \$$\frac{1}{8}$ to \$$\frac{1}{16}$) and January 2001 (from \$$\frac{1}{16}$ to \$0.01).[17] Chordia, Roll, and Subrahmanyam (2007) find a similar pattern when they use the quoted and effective spreads. Because changes in tick sizes are unlikely to have a material effect on order-processing and inventory-holding costs, we conjecture that the smaller non-information components under smaller tick sizes may largely be due to reduced market-maker rents.

---

[15] See Hasbrouck (1999, 2005, and 2009) and Chordia, Huh, and Subrahmanyam (2009).
[16] The results are qualitatively similar when variables are not scaled by price.
[17] See Chakravarty, Harris, and Wood (2001) and the SEC's (2012) *Report to Congress on Decimalization* for details on the minimum tick size changes in the U.S. securities markets.

When we separate out the transitory variable component from the total transitory costs based on the Sadka (2006) model, we find (as shown in Figure 1) that the transitory variable component $(\bar{\lambda}^S)$ is negative on average and close to zero with no discernible trend or variation over time, suggesting that the transitory variable cost is likely to be unimportant in asset pricing.

The strong time-varying nature of the components in trading costs shown in Figure 1 suggests that the role and relative importance of the cost components in asset pricing may vary across different time periods and stocks. Therefore, it would be difficult to draw reliable and general inferences as to the impact of these costs on expected stock returns from an analysis that uses an old database (ISSM), a small subset of stocks, or a short study period (e.g., Brennan and Subrahmanyam, 1996; Huang and Stoll, 1997). This is especially the case given the significant changes in market structure, trading technologies, and trading behaviors during the last two decades. Our study uses a broad and long dataset processed via the two different matching methods depending on the level of HFT volume.

IV. The Role of Non-Information Costs of Trading in Asset Pricing

A. Methodology and Variable Construction

Following Brennan, Chordia, and Subrahmanyam (1998), Ang, Liu, and Schwarz (2008), and Chordia, Huh, and Subrahmanyam (2009), we conduct our asset-pricing tests using data on *individual* securities.[18] Chordia, Subrahmanyam, and Tong (2014) use risk-adjusted returns in asset-pricing regressions. Similarly, Brennan, Huh, and Subrahmanyam (2015a) use risk-adjusted returns to examine the asymmetric effects of informed trading on stock returns. Following these studies, we calculate *risk-adjusted* returns. Our goal in this section is to investigate whether the non-information component of trading costs has any incremental explanatory power for stock returns beyond what the Fama-French (FF, 1993) three factors (FF3: MKT_t, SMB_t, and HML_t), Carhart's (1997) momentum factor (UMD_t), Pastor-Stambaugh's

[18] See Ang et al. (2008) for the advantages of using individual stocks (instead of portfolios) in asset-pricing tests.

(2003) liquidity factor (LIQ_t), and other firm attributes predict. Following the convention described above, we estimate the five-factor (F5)-adjusted returns in two ways.[19] In the first method, we calculate the F5-adjusted return (R_{jt}^{e1}) in each month using the following formula:

$$R_{jt}^{e1} = \left(\tilde{R}_{jt} - R_{Ft}\right) - \left(\hat{\beta}_{j1}^{*}MKT_t + \hat{\beta}_{j2}^{*}SMB_t + \hat{\beta}_{j3}^{*}HML_t + \hat{\beta}_{j4}^{*}UMD_t + \hat{\beta}_{j5}^{*}LIQ_t\right). \qquad (9)$$

To calculate the F5-adjusted return above, we estimate the factor loadings for each stock using the entire sample range of the data (from January 1983 to December 2010). In this method, therefore, we have only one set of the five factor loadings ($\hat{\beta}_{jk}^{*}$) for each stock.

To consider the time-varying features of factor betas, in the second method, we obtain the rolling estimates of the factor loadings, β_{jk} (k =1 to 5), in each month using the time series of the past 60 months (at least 24 months). Given the current month's data ($\tilde{R}_{jt} - R_{Ft}$, MKT_t, SMB_t, HML_t, UMD_t, LIQ_t) and the estimated factor loadings ($\hat{\beta}_{jk}^{**}$), we then calculate the second F5-adjusted return, R_{jt}^{e2}, in each month using the following formula:

$$R_{jt}^{e2} = \left(\tilde{R}_{jt} - R_{Ft}\right) - \left(\hat{\beta}_{j1}^{**}MKT_t + \hat{\beta}_{j2}^{**}SMB_t + \hat{\beta}_{j3}^{**}HML_t + \hat{\beta}_{j4}^{**}UMD_t + \hat{\beta}_{j5}^{**}LIQ_t\right). \qquad (10)$$

Using one of the two types of the F5-adjusted return obtained from Eq. (9) and Eq. (10), we conduct monthly Fama-MacBeth (1973) cross-sectional regressions as follows:

$$R_{jt}^{eh} = c_{0t-1} + \gamma\Lambda_{jt-1}^{i} + \textstyle\sum_{n=1}^{N} c_{nt}\mathrm{X}_{njt-1} + \tilde{e}_{jt}'; \qquad (11)$$

where h = 1 or 2, Λ_{jt-1}^{i} is the non-information component of trading costs, and X_{njt-1} denotes attributes (n = 1, …, N) of firm j in month t – 1 [firm attributes (*SIZE*, *BTM*, and *PAR1-PAR4*) will be explained below]. For comparison purposes, we also estimate Eq. (11) using the F5-*unadjusted* excess returns (R^e). The coefficient vector $\boldsymbol{c}_t = [c_{0t}\ \gamma_t\ c_{1t}\ c_{2t} \dots c_{Nt}]'$ in Eq. (11) is estimated each month with ordinary least-squares (OLS) regressions. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and its standard error is taken from the time series of the monthly coefficient estimates.[20]

---

[19] The reason that we use the F5-adjusted return as the dependent variable in Eq. (11) is to avoid the errors-in-variables (EIV) problem. We may estimate the five factor betas and directly include them (for X_{t-1}) as control variables in Eq. (11). In that case, however, we cannot get around the EIV problem. That is why we adjust the dependent variable in two ways and use one of two F5-adjusted returns in Eq. (11).

[20] We construct all the explanatory variables in Eq. (11) with past data relative to the dependent variable, which is based on the last trading day's closing price within each month. For example, we estimate Λ_t^i with intraday order flows from the beginning of

Avramov and Chordia (2006) suggest that the constant-beta version of the Fama and French (1993) three-factor model does not adequately capture the effect of firm attributes on stock returns. We thus consider the following firm attributes in our analysis: *SIZE* ≡ the natural logarithm of the market value of equity (*MV*); *BM* ≡ the book-to-market ratio (BV/MV), where BV and MV are the book and market values of equity in million dollars;[21] BTM ≡ the Winsorized value (at the 0.5th and 99.5th percentiles) of *BM;* and *PAR1*, *PAR2*, *PAR3*, and *PAR4* ≡ the compounded holding period returns over the most recent three months (from month *t*-1 to month *t*-3), from month *t*-4 to month *t*-6, from month *t*-7 to month *t*-9, and from month *t*-10 to month *t*-12, respectively. For each of the momentum variables to exist, a stock must have all three returns over the corresponding three-month period. We calculate the book-to-market ratio using data from the CRSP and CRSP/Compustat Merged (CCM) files, and calculate other variables using data from the CRSP monthly file.

Panel E in Table 2 shows the time-series average of monthly cross-sectional means, medians, standard deviations (*STD*), and other descriptive statistics for each firm attribute. The average market value of equity (*MV*) is \$3.39 billion and the average book-to-market ratio (*BM*) is 0.69 over the sample period. Both variables are highly skewed and leptokurtic, but their adjusted values (*SIZE* and BTM) are not. Hence, we use *SIZE* and *BTM* (instead of *MV* and *BM*) as control variables in our asset-pricing tests. The average values of the momentum variables ($PAR1$-$PAR4$) increase monotonically, ranging from 1.3% to 1.7%.

We report in Table 3 the correlations (the time-series average of monthly cross-sectional correlation coefficients) between the key variables of interest in our study. The three non-information cost measures $\left(\frac{\bar{\varphi}^{GH}}{P}, \frac{\bar{\varphi}^{FV}}{P}, \text{and } \frac{\bar{\varphi}^{S}}{P}\right)$ are highly correlated with each other. Given that the third measure $\left(\frac{\bar{\varphi}^{S}}{P}\right)$ is separated out from the transitory variable component, it is correlated to

the first trading day to the end of the last trading day within a month, and then scale it by the previous month-end stock price. BTM is six-month lagged as defined below; *SIZE* is based on the market value as of the previous month end; and the momentum variables ($PAR1$-$PAR4$) are the returns of past months. Therefore, we use the time subscript "t – 1" for the explanatory variables.

[21] In line with Fama and French (1992), we lag the quarterly book-to-market ratio by two quarters (assuming a lag of six months before the data are known to investors). We then convert the quarterly data into a monthly series.

a lesser degree with the first two measures. When the transitory costs are decomposed into the fixed and variable components ($\frac{\bar{\varphi}^S}{P}$ and $\frac{\bar{\lambda}^S}{P}$), the two are negatively correlated. The first three (price-scaled) non-information cost measures are negatively correlated with firm size (*SIZE*), which implies that a larger firm faces a smaller non-information trading cost per one-dollar trading. The non-information cost measures (except for the fourth) are also negatively correlated with the momentum variables (*PAR1-PAR4*). This result suggests that better price performance in the past improves the liquidity of a stock during the current period, which is consistent with Lee and Swaminathan (2000).

B. Fama-MacBeth Regression Results

We conduct cross-sectional regressions in a multivariate setting to investigate whether the non-information component of trading costs has any impact on stock returns. Table 4 reports the results of the Fama-MacBeth regressions specified in Eq. (11). In addition to the average coefficients and *t*-statistics,[22] we also report the average of the adjusted R^2 values from the individual regressions (denoted by *Avg R-sqr*) and the mean number of stocks (*Avg Obs*) used in the regressions each month over the sample period. Panel A in Table 4 shows the results using the non-information cost ($\frac{\bar{\varphi}^{GH}}{P}$) estimated by the Glosten and Harris (1998) model, while Panel B shows the results using the non-information cost ($\frac{\bar{\varphi}^{FV}}{P}$) estimated by the Foster and Viswanathan (1993) model. For expositional convenience, we multiply all the regression coefficients by 100 in Table 4 and subsequent asset-pricing test tables, since the coefficients on some variables are small.

The results show that the explanatory power of the model is higher with the unadjusted excess returns (R^e) than with the F5-adjusted returns (R^{e1} and R^{e2}), indicating that the Fama-French factors and the other two factors have some ability to price stocks in the cross-section.

[22] We find no evidence of significant autocorrelations in the time series of the estimated coefficients (the absolute values of the first-order serial correlations in the coefficient series are lower than 10%). Therefore, we report the standard Fama-MacBeth *t*-statistic instead of the Newey and West (NW) (1987, 1994) *t*-statistic throughout the paper.

Because estimation of R^{e2} requires stricter conditions (e.g., return series must exist for at least 24 months within the past 60-month window), the average number (1,686) of component stocks is smaller when the dependent variable is R^{e2} than the corresponding figure (1,844) when it is either R^{e} or R^{e1}.

Panel A in Table 4 shows that the average coefficients on the non-information cost measure ($\frac{\bar{\varphi}^{GH}}{P}$) estimated from the Glosten and Harris (1998) model are positive and significant at the 1% level, regardless of whether we use the unadjusted excess return (R^{e}) or the F5-adjusted excess returns (R^{e1} and R^{e2}) as the dependent variable. Likewise, Panel B shows that the average coefficients on the non-information cost measure ($\frac{\bar{\varphi}^{FV}}{P}$) estimated via the Foster and Viswanathan (1993) model are positive and significant in all three regressions. We note that the estimated coefficients on $\frac{\bar{\varphi}^{FV}}{P}$ and their *t*-statistics are slightly smaller than the corresponding figures for $\frac{\bar{\varphi}^{GH}}{P}$.[23] This result may be due to the fact that $\frac{\bar{\varphi}^{FV}}{P}$ is based only on unexpected order flows. On the whole, these results indicate that firms with a higher non-information cost of trading exhibit a higher equilibrium stock return in the U.S. equity market.

Panels A and B also shows that none of the coefficients on firm size (*SIZE*) is significantly different from zero in all six regressions. The regression coefficients on the book-to-market ratio (*BTM*) are significant and positive when we use the unadjusted returns, but they tend to become insignificant with the F5-adjusted returns. We also find that the coefficients on *PAR2-PAR4* are all positive and significant, which is consistent with the finding of prior research (Jegadeesh and Titman, 1993).

Panel C in Table 4 reports the regression results when we decompose the non-information cost into the transitory fixed cost ($\frac{\bar{\varphi}^{S}}{P}$) and the transitory variable cost ($\frac{\bar{\lambda}^{S}}{P}$) based on Sadka (2006). When we include only the transitory fixed cost in the regression (i.e., columns 1, 4, and 7), the coefficients on $\frac{\bar{\varphi}^{S}}{P}$ are all positive and significant. Sadka (2006) reports that the

[23] The *t*-statistics for the coefficients on the non-information component in Table 4 (and other subsequent asset-pricing results) are well above the critical threshold of 3.0 recommended by Harvey, Liu, and Zhu (2013). See their paper for a caution against over-emphasizing the variables that are only marginally significant in asset-pricing tests.

systematic liquidity risk associated with the transitory fixed component does not explain portfolio returns. These results suggest that it is the transitory fixed cost (*level*) itself, rather than its systematic *risk* (measured by beta), that commands a positive return premium in the cross-section of stocks. Note in specification 7 in Panel C that a one-standard-deviation increase in the transitory fixed cost ($\frac{\bar{\varphi}^S}{P}$) is associated with a 0.93% increase (per month) in the F5-adjusted excess return computed by the second method (R^{e2}).

When we include only the transitory variable cost in the regression (i.e., see specifications 2, 5, and 8 in Panel C), the coefficients on $\frac{\bar{\lambda}^S}{P}$ are negative and tend to be significant. However, when we include the transitory fixed and variable costs together (i.e., see specifications 3, 6, and 9), the coefficients on $\frac{\bar{\lambda}^S}{P}$ mostly become insignificant, but the coefficients on $\frac{\bar{\varphi}^S}{P}$ remain positive and significant. These results indicate that the fixed cost of market making (e.g., the order processing cost and the inventory holding cost) is priced, while the variable cost of market making (which increases with the size of order flows) is less important in asset pricing. This finding also suggests that decomposing the non-information cost into the transitory fixed component ($\frac{\bar{\varphi}^S}{P}$) and the transitory variable component ($\frac{\bar{\lambda}^S}{P}$), as in Sadka (2006), might not be necessary in asset-pricing studies.[24] In subsequent analysis, therefore, we focus only on the non-information cost measures estimated by the Glosten-Harris (1988) and Foster-Viswanathan (1993) models.

Finally, Panel C shows that when the two non-information cost elements [that are decomposed based on Sadka (2006)] are employed in the cross-sectional regressions, the size effect becomes statistically significant with the risk-adjusted returns used as the dependent variable. The book-to-market effect is again weakened when R^{e2} is used as the dependent variable. The momentum effects are similar to those shown in Panels A and B.

In summary, as the magnitude of the coefficients and their *t*-values suggest, the impact of

[24] However, when the adverse-selection cost is decomposed into the permanent fixed cost ($\frac{\varphi^S}{P}$) and the permanent variable cost ($\frac{\lambda^S}{P}$), as in Sadka (2006), we find (in the unreported result) that the permanent fixed cost ($\frac{\varphi^S}{P}$) also plays a strong role in asset-pricing regressions.

the non-information cost on stock returns is so strong that its role in asset pricing should not be ignored. The next questions are then whether the effects of the non-information component on returns are robust in different settings, and how the size and role of the non-information component compare to those of the adverse-selection component. We address these questions in Section VI.

V. Robustness Tests

A. Results with the Sample Excluding the ISSM or High-Frequency-Trading Years

A potential issue in using high-frequency data is the accuracy of ISSM. For instance, in the early years of ISSM, the data were entered by hand, which could have led to some errors or missing observations. In addition, in the TAQ database, some condition codes for identifying different types of trades are not exactly the same as those in ISSM. To address these issues, we replicate our empirical analysis reported in Section IV using only the TAQ data (1993-2010: 18 years). Dropping the 10-year period (1983-1992) also seems logical as a robustness check because the financial market environments may have changed since the early 1990s.

We report the results in Panel A and Panel B of Table 5. By precluding the 10 years of the ISSM period, the average number of component stocks increases by more than 12%, compared to that reported in Table 4. In addition, we find that the momentum effects are weaker, relative to the results in Table 4. In contrast, the size effect becomes positive and significant. Of greater interest is that the impact of the two non-information cost measures (i.e., the size of the coefficients) on returns becomes even larger. Although the *t*-values for the non-information component are smaller than those in Table 4, they are all greater than 7, indicating the important role of the non-information cost in asset pricing during this shorter, more recent time period.

O'Hara, Yao, and Ye (2011) show that missing odd-lot volume in TAQ ranges from 2.25% to 4.0% in recent years when HFT volume (and hence odd-lot trading) is large. Prior research (Chakrabarty, Moulton, and Shkilko, 2012; Easley, Lopez de Prado, and O'Hara, 2012) also shows that the accuracy of the Lee and Ready (1991) method is lower during the HFT era.

Because we use the Holden-Jacobsen algorithm for the HFT years (2007-2010) to minimize the potential misclassification of trades, our main analyses shown in Table 4 may not be subject to significant errors. Nonetheless, for consistency in using the Lee-Ready method, we replicate our analysis using the sample up to 2006 only, excluding the last four HFT years (2007-2010).

The results reported in Panel C and Panel D of Table 5 show that the size and significance of the coefficients on the two non-information cost measures are qualitatively identical to those reported in Table 4. The book-to-market and momentum effects as well as the (negative) size effect all tend to become stronger after we drop the last four years from the sample. This result is consistent with the finding by Chordia, Subrahmanyam, and Tong (2014) that the return predictability of firm attributes has decreased substantially in recent years.

B. Controlling for Idiosyncratic Volatility and Earnings Surprise

Recent studies have identified additional firm characteristics and anomalies that predict stock returns. Notably, Chordia, Subrahmanyam, and Tong (2014) show that idiosyncratic volatility (*IVOL*) and earnings surprise (*SUE*) predict stock returns, although improved liquidity and increased arbitrages have decreased their predictive power over time. In this subsection, we conduct our analysis with these additional control variables. *IVOL* is the standard deviation of residuals from the time-series regression of the monthly excess return of each stock on the five factors (MKT_t, SMB_t, HML_t, UMD_t, and LIQ_t) using the rolling window of the past 60 months. As in Livnat and Mendenhall (2006), we measure earnings surprise by $SUE_{it} = \frac{EPS_{it} - EPS_{it-4}}{P_{it}}$, where EPS_{it} is the "street" earnings per share for firm *i* in quarter *t* that excludes special items from the Compustat-reported EPS [following Abarbanell and Lehavy (2007)]; P_{it} is the stock price at the end of quarter *t*; and EPS_{it-4} is the EPS at the end of quarter *t-4* (adjusted for stock splits and stock dividends). The advantage of this *SUE* definition is that quarterly earnings surprise can be estimated as long as EPS data are available, unlike other *SUE* definitions that require analysts' forecasts.

We report the regression results in Table 6. The sample sizes in Table 6 are smaller than those in Table 4 because of missing *SUE* values for some firms. As both panels in Table 6 show, our results are robust to controlling for *IVOL* and *SUE*. The impact of the non-information cost on stock returns continues to be strong, regardless of which model is used for estimating the cost component. Consistent with the finding of prior studies, we also find that *IVOL* is negatively related to stock returns (Ang et al., 2006), while *SUE* is positively related to stock returns (Bernard and Thomas, 1989).

C. Under Different Tick-Size Regimes

The New York Stock Exchange (NYSE) reduced the minimum tick size from $\$\frac{1}{8}$ to $\$\frac{1}{16}$ on June 24, 1997 and decimalized it in January 2001. Given that our non-information cost measures in Figure 1 reflect these changes sharply, we examine whether the impact of the non-information cost on stock returns varies across different tick-size regimes. The two sub-periods that we consider are the $\$\frac{1}{8}$ era (January 1983 to May 1997) and the decimal era (February 2001 to December 2010). The middle period ($\$\frac{1}{16}$ regime) from July 1997 to December 2000 is not considered because the interval is too short to calculate the Fama-MacBeth statistics. For this experiment, we exclude AMEX-listed stocks from the study sample.[25] To save space, we report in Table 7 only the results using the Foster and Viswanathan (1993) measure, which adjust for serial dependence in order flows.

The results show that the coefficients on $\frac{\bar{\varphi}^{FV}}{P}$ are positive and significant in both sub-periods, regardless of how the returns are adjusted with regard to the five factors. Considering that the decimal era is relatively short (about 10 years) for computing the Fama-MacBeth statistics, our results again underscore the critical role of the non-information cost in asset pricing.

Comparing the results in the two panels of Table 7 with those reported in Panel B of

[25] The AMEX began to use the $\$\frac{1}{16}$ tick size from 1992 for stocks priced between \$0.25 and \$5 and extended in 1997 to all stocks trading at or above \$0.25 [see U.S. SEC's (2012) report]. To avoid any confounding effect, therefore, we exclude AMEX-listed stocks in this test.

Table 4, the coefficients on $\frac{\bar{\varphi}^{FV}}{P}$ are smaller in the $\$\frac{1}{8}$ era but larger in the decimal era than the corresponding coefficients for the entire study period (1983-2010). This may be due to the difference in the level of $\frac{\bar{\varphi}^{FV}}{P}$ across these periods, as Figure 1 suggests. That is, all else being equal, the smaller the value of an independent variable, the larger the magnitude of the coefficient on that variable. We also find that the momentum effects (see *PAR2-PAR4*) are strong in the $\$\frac{1}{8}$ era, but not likely to be significant in the decimal era. The coefficients on *SIZE* and *BTM* have incorrect signs in the decimal era, perhaps because of the short estimation period or the reasons described in Chordia et al. (2014).

D. Weighted Least-Squares (WLS) Regressions

Asset-pricing tests using CRSP returns may suffer from a microstructure bias that results from the bid-ask bounce. Asparouhova, Bessembinder, and Kalcheva (2010) suggest that weighted least-squares (WLS) estimation can reduce this problem. Following their approach, we conduct WLS regressions using the prior-month gross return (one plus the return in month *t*-1) as a weighting variable. The results (not reported for brevity) show that although the coefficients on the two non-information cost measures ($\frac{\bar{\varphi}^{GH}}{P}$ and $\frac{\bar{\varphi}^{FV}}{P}$) and their *t*-values are somewhat smaller than the corresponding values reported in Table 4, the sign and significance of the coefficients remain qualitatively identical.

VI. Relative Importance of the Adverse-Selection and Non-Information Costs

A. Relative Size of the Two Components

To assess the magnitude of the non-information cost of trading relative to the adverse-selection cost, we calculate the proportion of the non-information component ($ProNIC^i$) and the proportion of the adverse-selection component ($ProASC^i$) for a trade of \$10,000 using the following formula:[26]

$$ProNIC^i = \frac{\bar{\varphi}^i}{\bar{\varphi}^i + \lambda^i}\text{ , } i = \text{GH or FV}, \tag{12}$$

$$ProASC^i = \frac{\lambda^i}{\bar{\varphi}^i + \lambda^i}\text{ , } i = \text{GH or FV}, \tag{13}$$

where $\bar{\varphi}^i$ and λ^i are the non-information cost and the adverse-selection cost, respectively, estimated by the Glosten and Harris (1988) model as specified in Eq. (5) or by the Foster and Viswanathan (1993) model as specified in Eq. (7).

We report in Table 8 the proportion of each component across the portfolios formed by firm size (*MV*) as well as for the whole sample. The proportion of the adverse-selection component (*ProASC*) is about 30% in the smallest-size quintile (*MV1*), regardless of whether we use raw order flows (Glosten and Harris, 1988) or unexpected order flows (Foster and Viswanathan, 1993) in the estimation. As firm size increases, *ProASC* decreases monotonically, with the fraction being only 2% in the largest-size group (*MV5*). For the whole sample, the proportion of the adverse-selection component (*ProASC*) accounts for about 12%. These results suggest that the adverse-selection cost is relatively high in small stocks, but on average the non-information component (*ProNIC*) is indeed dominant (88%) in stock trading costs, confirming the small sample result of Huang and Stoll (1997).[27]

B. Relative Importance of the Two Components in Asset Pricing

B.1. Controlling for the Adverse-Selection Component

[26] Because $\bar{\varphi}$ is multiplied by 10^2 and λ by 10^6 in our tables, their original values ($\bar{\varphi}^*$ and λ^*) from Eqs. (5) and (7) could be expressed as $\bar{\varphi}^* = (10^{-2})\bar{\varphi}$ and $\lambda^* = (10^{-6})\lambda$. In the spirit of Glosten and Harris (1988) (see pp. 135-136 and Table 2), the spread proxy for a round-trip transaction of V dollars is given by $2[\bar{\varphi}^* + (\lambda^*)(V)] = 2(10^{-2})[\bar{\varphi} + \lambda(10^{-4})V] = \frac{2}{100}[\bar{\varphi} + \lambda(10^{-4})V]$. Thus, for a transaction of \$10,000 (i.e., V = 10,000), the spread proxy becomes $\frac{2}{100}[\bar{\varphi} + \lambda]$. Because this spread proxy is proportional to $(\bar{\varphi} + \lambda)$, we use $SUM^i = (\bar{\varphi}^i + \lambda^i)$ (i = *GH* or *FV*) as a cost proxy equivalent to the spread (ignoring 2/100). Then, its price-scaled version is $SUM_P^i = \left(\frac{\bar{\varphi}^i + \lambda^i}{P}\right)$ (i = *GH* or *FV*). Also, the proportion of the non-information component is given by $\frac{\bar{\varphi}/P}{(\bar{\varphi}+\lambda)/P} = \frac{\bar{\varphi}}{\bar{\varphi}+\lambda}$. Therefore, we denote the proportions of the non-information component and the adverse-selection component by $ProNIC^i = \frac{\bar{\varphi}^i}{\bar{\varphi}^i+\lambda^i}$ and $ProASC^i = \frac{\lambda^i}{\bar{\varphi}^i+\lambda^i}$ (i = *GH* or *FV*), respectively, for a stock transaction of \$10,000.

[27] Huang and Stoll (1997) estimate spread components for a sample of 20 stocks using one-year (1992) data from ISSM and report that the non-information component is larger than the adverse-selection component. However, Huang and Stoll (1997) do not examine the issue of whether the non-informational component is priced or not in the cross-section of stock returns.

To evaluate the relative importance of the adverse-selection and non-information components of trading in the asset-pricing context, we consider both the adverse-selection cost measures ($\frac{\lambda^{GH}}{P}$ or $\frac{\lambda^{FV}}{P}$) and the non-information cost measures ($\frac{\bar{\varphi}^{GH}}{P}$ or $\frac{\bar{\varphi}^{FV}}{P}$) in the cross-sectional regressions. Panel A in Table 9 reports the regression results when we measure the two components using the Glosten and Harris (1988) model and Panel B reports the results based on the Foster and Viswanathan (1993) model.

The first three columns in each panel show the results with R^{e1} as the dependent variable and the next three columns show the results with R^{e2}. For comparison purposes, specifications 1 and 4 in each panel reproduce the regression results for the non-information component that we reported earlier in Table 4. Specifications 2 and 5 show the results when we include only the adverse-selection component, and specifications 3 and 6 show the results when we include both the non-information and adverse-selection components, together with the control variables.

The results (see regressions 2 and 5 in Panels A and B) show that the F5-adjusted returns are positively and significantly related to the adverse-selection cost ($\frac{\lambda^{GH}}{P}$ or $\frac{\lambda^{FV}}{P}$) after accounting for the effects of firm characteristics, which is consistent with Chordia, Huh, and Subrahmanyam (2009). When both components are included in the same equation (regressions 3 and 6 in each panel), we find that the coefficients on the adverse-selection component ($\frac{\lambda^{GH}}{P}$ or $\frac{\lambda^{FV}}{P}$) become smaller by 56%-74%, although they are statistically significant. By contrast, the coefficients on the non-information component ($\frac{\bar{\varphi}^{GH}}{P}$ or $\frac{\bar{\varphi}^{FV}}{P}$) continue to be strongly significant, with their *t*-values being two-digit numbers when R^{e1} is used. Other interesting aspect is that the size and book-to-market effects are often weakened when we include the non-information component, whereas those effects tend to remain strong when we include only the adverse-selection component in the regression (specifications 2 and 5).

The average standard deviation of $\frac{\lambda^{GH}}{P}$ and $\frac{\lambda^{FV}}{P}$ is 0.568 and 0.576, respectively. Using this information (and Panel D in Table 2), we find in specification 1 of Panel B that a one-standard-deviation increase in $\frac{\bar{\varphi}^{FV}}{P}$ causes excess returns (R^{e1}) to increase by 1.32% (= 2.778*0.474) per month, while the corresponding figure for $\frac{\lambda^{FV}}{P}$ in specification 2 is only 0.67%

(= 1.163*0.576). More importantly, in specification 3 of Panel B, the impact of the former is 1.37%, but that of the latter is only 0.19%. We obtain similar results with R^{e2} or with the cost measure estimated with the other model. These results indicate that the non-information component plays a more important role in asset pricing than does the adverse-selection component, regardless of how returns are risk-adjusted or how cost components are estimated.

B.2. Standardized Coefficients

We estimate standardized coefficients to shed further light on the relative importance of the non-information and adverse-selection components. Standardized coefficients allow us to measure changes in stock returns per one-standard-deviation increase in the explanatory variables. We estimate monthly standardized coefficients and tabulate the Fama-MacBeth and other statistics in Panel of A and Panel B of Table 10. To save space, in this and next subsections we report only the results from regressions that use R^{e2}, which incorporates the time-varying features of five factor betas. Specification 4a in Panel A shows that the impact of a one-standard-deviation increase in the non-information component ($\frac{\bar{\varphi}^{GH}}{P}$) on returns is 7.182. The corresponding figure for the adverse-selection component ($\frac{\lambda^{GH}}{P}$) is only 3.633 (see specification 5a). Similarly, specification 6a contrasts the effects of $\frac{\bar{\varphi}^{GH}}{P}$ on stock returns with that of $\frac{\lambda^{GH}}{P}$: 7.356 vs. 0.881. When the cost components are estimated based on Foster and Viswanathan (1993) in Panel B, the patterns are virtually the same. Overall, these results are qualitatively similar to those reported above.

B.3. Trade-Size Adjustments for the Adverse-Selection Component

We show in Table 9 that the relation between the F5-adjusted return and the adverse-selection component becomes weaker when we include the non-information component in the regression. It is possible that this result could result from a high correlation between the two components. [For our sample stocks, the average correlation between $\frac{\lambda^{GH}}{P}$ and $\frac{\bar{\varphi}^{GH}}{P}$ is 0.533 and that between $\frac{\lambda^{FV}}{P}$ and $\frac{\bar{\varphi}^{FV}}{P}$ is 0.522.] Brennan and Subrahmanyam (1996) show that when the adverse-selection component is adjusted for trade size, it is less correlated with the non-information component. To examine whether our results in Table 9 are driven by the correlation between the two components, we replace $\frac{\lambda^{GH}}{P}$ and $\frac{\lambda^{FV}}{P}$ with their trade-size-adjusted equivalents C_q_GH and C_q_FV, where $C_q\_GH = \left(\frac{\lambda^{GH}}{P} * q\right)$, $C_q\_FV = \left(\frac{\lambda^{FV}}{P} * q\right)$, and q is the average trade size within each month (in \$100).[28] We calculate q for each firm by dividing the total dollar trading volume within each month (in \$100) by the total number of trades within each month (i.e., the number of buyer-initiated trades + the number of seller-initiated trades within each month). The regression results (see Panel C and Panel D in Table 10) show that although *t*-values for the coefficients on the trade-size-adjusted adverse-selection components are larger than the corresponding values in Table 9, the estimated coefficients on the two components are qualitatively similar to those reported in Table 9.

To further address the potential multicollinearity problem, we also calculate the time-series average of tolerance values (*AvgTol*) and variance inflation factors (*AvgVIF*) obtained from monthly regressions over the whole sample period (1983:01-2010:12, 336 months) for two regression specifications (regressions 3 and 6) reported in Panel B of Table 9. The results (not tabulated) show that our regressions are not subject to the multicollinearity problem.[29]

On the whole, our empirical results underscore the relative importance of the non-information cost of trading in asset pricing. A number of prior studies have focused on the asset-pricing implication of the adverse-selection component of trading costs. However, our results show that it is the non-information component that plays a more important role in asset pricing when both components are considered together. There are several plausible explanations for why

[28] Consistent with the finding of Brennan and Subrahmanyam (1996), we find that the non-information component is less correlated with the trade-size-adjusted adverse-selection component than with the unadjusted adverse-selection component. The mean correlation coefficient between C_q_GH and $\frac{\overline{\varphi}^{GH}}{P}$ is 0.389 and the mean correlation coefficient between C_q_FV and $\frac{\overline{\varphi}^{FV}}{P}$ is 0.377.

[29] The test results show that, for all explanatory variables in each specification, *AvgTol* is larger than 0.10, which is the minimum acceptable value in assessing the multicollinearity problem. Similarly, in each specification, the largest *AvgVIF* is 2.08, which is much smaller than the maximum acceptable value (10). Therefore, even with (much stricter) 5 as the critical value for the variance inflation factor, our regressions do not suffer from the multicollinearity problem.

the non-information component plays a more prominent role in asset pricing than the adverse-selection component does. First, the non-information component may play a stronger role simply because investors pay more attention to it, given its sheer size relative to the adverse-selection component. The adverse-selection component may also indirectly capture the non-information component and the former may lose a portion of its explanatory power once the latter is included in the regression, because of a stronger link of the non-information component to stock returns.

Another possible explanation is that investors may be able to more readily reflect the non-information cost in their required returns because it is easier for them to measure the non-information component. Assessing the adverse-selection cost is likely to be inherently difficult because detecting the presence of information asymmetry and measuring its size accurately may require significant expertise and efforts. In contrast, it may be easier for investors to measure the non-information component because it is more closely related to observable stock attributes such as trading volume. The non-information component may be easier to estimate also because it depends on the variables, such as minimum tick sizes, that vary across stocks but do not vary across trades in a given stock. Consequently, investors' perception and measurement of the adverse-selection component may contain more noise than those of the non-information component.

In support of the above conjecture, we report in Table 11 the average *t*-values and the proportions of positive and significant cost elements, which are estimated from the time-series regressions as specified in Eqs. (5) and (7). As shown in the table, the estimates of the non-information component are much more accurate (i.e., smaller standard errors and larger *t*-values) on average than the estimates of the adverse-selection component.[30] Specifically, more than

[30] Prior research provides mixed results regarding the relation between stock returns and the bid-ask spread. Amihud and Mendelson (1986) provide initial evidence that stock returns are positively related to the bid-ask spread. However, many subsequent studies show that stock returns do not increase with the bid-ask spread in the U.S. and international stock markets. For instance, Easley et al. (2002) and Chordia et al. (2009) report a negative relation between the spread and the return in the U.S. stock market. Han et al. (2012) also report a negative or no relation in international markets (45 countries including the U.S.). Hasbrouck (2009) and Chordia et al. (2009) show that Roll (1984)'s spread measure is not priced, either. A potential explanation of the mixed evidence is that the bid-ask spread is a noisy measure of trading costs because it contains a very noisy component (i.e., the adverse-selection component).

99% of the estimates of the non-information component ($\bar{\varphi}_0^{GH}$ and $\bar{\varphi}_0^{FV}$) are positive and significant at the 5% level and the pooled sample mean of individual *t*-values is as large as 54. In contrast, only 81%-82% of the estimates of the adverse-selection component (λ_0^{GH} and λ_0^{FV}) are positive and significant, and the mean of *t*-values is much smaller than that for the non-information component. Overall, these results support the idea that the measurement of the non-information component involves smaller errors.

VII. Additional Tests Using the Probability of Informed Trading and the Amihud Measure

Easley, Hvidkjaer, and O'Hara (2002) and Brennan, Huh, and Subrahmanyam (2015a) report that stock returns are positively associated with the probability of informed trading (*PIN*). Brennan, Huh, and Subrahmanyam (2015a) decompose *PIN* into two components [good-news *PIN* (*PIN_G*) and bad-news *PIN* (*PIN_B*)] and show that the bad-new *PIN*, *PIN_B*, drives the pricing of *PIN*.

Several recent studies suggest that information asymmetry alone is unlikely to explain the cross-sectional variation in stock returns or the cost of capital. Notably, Hughes, Liu, and Liu (2007) show that for large economies, although greater information asymmetry leads to higher market-wide risk premiums, it does not affect the required rate of return in the cross-section because idiosyncratic risks could be eliminated through diversification.[31] Similarly, Lambert, Leuz, and Verrecchia (2012) show that under perfect competition, the required return is determined solely by the average precision of investors' information, not by the degree of information asymmetry. Armstrong et al. (2011) provide empirical evidence that is consistent with the prediction of Lambert, Leuz, and Verrecchia (2012).[32] Lambert and Verrecchia (2014)

[31] Hughes, Liu, and Liu (2007) analyze the relation between asymmetric information and the cost of capital (or expected stock returns) using the framework of a competitive noisy rational expectations economy (e.g., Grossman and Stiglitz, 1980; Admati, 1985; Easley and O'Hara, 2004). The authors point out that although numerous empirical studies cite Easley and O'Hara (2004) to invoke the notion that information asymmetry is priced because uninformed investors demand price protection from trading with privately informed investors, this notion is not supported by Easley and O'Hara (2004). They show that the pricing effect described in Easley and O'Hara's (2004) Proposition 2 is driven by under-diversification, which does not apply to large economies.

[32] They find evidence that the cost of capital is unrelated to information asymmetry in perfectly competitive markets.

confirm the single-asset result in Lambert, Leuz, and Verrecchia (2012) that under perfect competition information asymmetry is not a separately priced risk in a multi-asset setting. Lambert and Verrecchia (2014) also show that information asymmetry affects the required return through its association with market illiquidity under imperfect competition.

These studies imply that the non-information cost of trading (e.g., order-processing and inventory-holding cost) may play a more important role in asset pricing than the adverse-selection *risk*, because the risk associated with the former is less diversifiable than the latter. A large portion of the adverse-selection risk arises from firm-specific idiosyncratic shocks, which is diversifiable by investors and market makers.[33] In contrast, a large portion of the risk associated with inventory management tends to move together across stocks, because trading volume of individual stocks (which is an important determinant of the inventory-holding risk) tends to move together with trading volume of the market. For example, Chordia, Roll, and Subrahmanyam (2000) and Karolyi, Lee, and van Dijk (2012) find strong commonality in trading activities of individual stocks. Consequently, the *risk* associated with the non-information cost may not easily be diversified away, and thus the non-information cost could be more strongly priced in the cross-section of stock returns than the adverse-selection risk.

To examine whether (1) the positive relation between *PIN-B* and stock returns shown in Brennan et al. (2015a) remains significant when the non-informational cost is include in the regression *and* (2) the positive relation between the non-information cost and stock returns shown in our study remains significant when the adverse-selection risk (instead of the adverse-selection cost, $\frac{\lambda^{GH}}{P}$ or $\frac{\lambda^{FV}}{P}$) is included in the regression, we replicate the analysis reported in Table 10 using *PIN_B*. Following Brennan et al. (2015a), we estimate *PIN*-related parameters (θ) using the Yan and Zhang (2012) algorithm on a quarterly basis. This algorithm involves conducting the

[33] Competitive noisy rational expectations models cited above do not consider the role of market makers. These models consider information asymmetry among *investors* and its implications for asset prices. Hence, the inventory-holding risk from the perspective of dealers and its possible ramifications for asset pricing have not been analyzed in these models. In our analysis, we consider the role of traders/investors and *market makers* in the price discovery process to draw asset-pricing implications of the adverse-selection and non-information costs.

optimization procedure up to 125 times for each stock-quarter using 125 pre-specified sets of initial parameter values. Among the estimated candidates, the algorithm chooses the θ estimate [where $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ is a vector of the parameters required to compute *PIN* and its components] that minimizes corner solutions while maximizing the likelihood function $L(M|\theta) = \prod_{j=1}^{J} L(\theta|B_j, S_j)$, where M is the data set, B_j and S_j denote the number of buyer- and seller-initiated trades on day *j* (see the Appendix for details on estimating *PIN_B*). Although this procedure makes computations more difficult, Yan and Zhang (2012) show that the algorithm reduces the boundary-solution bias by effectively expanding the parameter space. Moreover, considering that informed trading can occur over a much shorter time interval than a year, our *quarterly* estimation of *PIN*-related parameters has an advantage over annual estimation because it allows for time-variation in the parameters.[34]

The estimation of *PIN_B* requires classifying each trade into a buyer- or seller-initiated category using the ISSM/TAQ databases and then counting the number of daily buys and sells (for data set M). Prior research (O'Hara, Yao, and Ye, 2011; Chakrabarty, Moulton, Shkilko, 2012) shows that *PIN* is not materially affected by the classification errors, owing to daily aggregation of buy and sell orders. To survive in our *PIN_B* estimation, stocks should have at least 40 positive-volume days within each quarter. For monthly regressions, the quarterly data are converted to a monthly series by filling the three months within each quarter with the corresponding quarterly estimate.

Panels A and B in Table 12 show the results when we replace the adverse-selection component with *PIN_B* in the regressions.[35] We report the standardized regression coefficients for the same reason described in Subsection B.2. The two panels show that the coefficients on the non-information cost measures ($\frac{\overline{\varphi}^{GH}}{P}$ and $\frac{\overline{\varphi}^{FV}}{P}$) are similar to those reported in Panel A and

[34] For example, Easley et al. (2002), Duarte and Young (2009), Mohanram and Rajgopal (2009), and Hwang et al. (2013) estimate *PIN* on an *annual* basis.

[35] The number of component stocks decreases by 5%-6% on average, compared to the sample size reported in Table 9. This means that the Yan and Zhang (2012) algorithm is still unable to calculate *PIN*-parameters for some stocks that trade very frequently.

Panel B of Table 10. We find that the effect of *PIN_B* is positive and significant. More importantly, the standardized coefficients on the non-information cost and their *t*-values are much larger than the coefficients on *PIN_B* and their *t*-values, indicating that stock returns are more strongly related to the non-information cost than to *PIN_B*.[36]

Brennan, Huh, and Subrahmanyam (2015a) suggest that the Amihud (2002) measure is more likely to capture the adverse-selection cost than to capture the non-information cost (see Table 2 of the paper). Brennan, Huh, and Subrahmanyam (2013) also show that the half-Amihud measure for down days (denoted by A^-) drives the pricing of the Amihud (2002) measure. As a further robustness check, we replace the adverse-selection component with the (turnover-version) half-Amihud measure, A^-, in our asset-pricing tests. We calculate A^- using the method in Brennan et al. (2013). As shown in Panel C and Panel D of Table 12, we again find the patterns in the regression coefficients similar to those in Table 10. That is, the standardized coefficients on the non-information cost measures and their *t*-values are much larger than the standardized coefficients on A^- and their *t*-values. As expected, the coefficients on A^- are positive and significantly different from zero. To sum up, all the above experiments together point to the notion that the non-information cost plays a more important role in asset pricing than the adverse-selection cost and risk.

VIII. Conclusion

In this study we examine the effect of the non-information cost of trading on stock returns. While numerous studies have analyzed the impact of information asymmetry on stock returns, little is known about the role of the non-information component in asset pricing. We provide strong evidence that the non-information component commands a positive return

[36] As shown in theoretical work by Hughes, Liu, and Liu (2007), information-asymmetry risks may be diversifiable and thus the non-information component of the spread plays a more important role in the cross-section of stock returns. In addition, information asymmetry may play a less important role in asset pricing because shares are traded in a sufficiently competitive environment (Lambert, Leuz, and Verrecchia, 2012; Lambert and Verrecchia, 2014).

premium after accounting for the effects of risk factors and firm characteristics. We conduct a variety of robustness tests and show that our main result remains qualitatively the same.

Our results also show that the non-information component is much larger than the adverse-selection component. More importantly, we find that the non-information cost of trading plays a more important role in asset pricing than does the adverse-selection cost or associated risk. We provide several plausible explanations for the stronger role of the non-information cost in asset pricing. Given the commonality in trading activities, the risk associated with the non-information cost may be more difficult to diversify away than the adverse-selection risk. Investors may pay more attention to the non-information component simply because it is larger. Finally, it may be easier for investors and traders to assess the non-information cost of trading and reflect it in asset prices than to assess the risk and cost of adverse selection, because the determinants of the former are more easily observable.

References

Abarbanell, J., and R. Lehavy, 2007, Letting the "tail wag the dog": the debate over GAAP versus street earnings revisited, Contemporary Accounting Research 24, 675-723.

Acharya, V., and L. Pedersen, 2005, Asset pricing with liquidity risk, Journal of Financial Economics 77, 375-410.

Admati, A., 1985, A noisy rational expectations equilibrium for multi-asset securities markets, Econometrica 53, 629-658.

Admati, A., and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, Review of Financial Studies 1, 3-40.

Amihud, Y., 2002, Illiquidity and stock returns: cross-section and time-series effects, Journal of Financial Markets 5, 31-56.

Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, Journal of Financial Economics 17, 223-249.

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

Ang, A., J. Liu, and K., Schwarz, 2008, Using stocks or portfolios in tests of factor models, Working paper, Columbia University.

Armstrong, C., J. Core, D. Taylor, and R. Verrecchia, 2011, When does information asymmetry affect the cost of capital? Journal of Accounting Research 49, 1-40.

Arnuk S. and J. Saluzzi, 2012, *Broken Markets*, FT Press/Upper Saddle River, NJ.

Atkins, A., and E. Dyl, 1997, Market structure and reported trading volume: NASDAQ versus the NYSE, Journal of Financial Research 20, 291-304.

Avramov, D. and T. Chordia, 2006, Asset pricing models and financial market anomalies, Review of Financial Studies 19, 1002-1040.

Asparouhova, E., H. Bessembinder, and I. Kalcheva, 2010, Liquidity biases in asset pricing tests, Journal of Financial Economics 96, 215-237.

Bernard and J. Thomas, 1989, Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1-36.

Bhattacharya, N., H. Desai, and K. Venkataraman, 2013, Does earnings quality affect information asymmetry? Evidence from trading costs, Contemporary Accounting Research 30, 482-516.

Bhattacharya, N., F. Ecker, P. Olsson, and K. Schipper, 2012, Direct and mediated associations among earnings quality, information asymmetry, and the cost of equity, Accounting Review 87, 449-482.

Brennan, M., T. Chordia, and A. Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, Journal of Financial Economics 49, 345-373.

Brennan, M., S. Huh, and A. Subrahmanyam, 2013, An analysis of the Amihud illiquidity premium, Review of Asset Pricing Studies 3, 133-176.

Brennan, M., S. Huh, and A. Subrahmanyam, 2015a, Asymmetric effects of informed trading on the cost of equity capital, Management Science, Forthcoming.

Brennan, M., S. Huh, and A. Subrahmanyam, 2015b, High-frequency measures of information risk, Working paper, UCLA.

Brennan, M., and A. Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, Journal of Financial Economics 41, 441-464.

Brunnermeier, M., and L. Pedersen, 2009, Market liquidity and funding liquidity, Review of Financial Studies 22, 2201-2238.

Campbell, J. Y., S. J. Grossman, and J. Wang, 1993, Trading volume and serial correlation in stock returns, Quarterly Journal of Economics 108, 905-39.

Carhart, M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57-82.

Chakrabarty, B., P. Moulton, and A. Shkilko, 2012, Short sales, long sales, and the Lee-Ready trade classification algorithm revisited, Journal of Financial Markets 15, 467-491.

Chakravarty, S., S. Harris, and R. Wood, 2001, Decimal trading and market impact, Working paper, University of Memphis.

Chordia, T., S. Huh, and A. Subrahmanyam, 2009, Theory-based illiquidity and asset pricing, Review of Financial Studies 22, 3629-3668.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, Commonality in liquidity, Journal of Financial Economics 56, 3-28.

Chordia, T., R. Roll, and A. Subrahmanyam, 2001, Market liquidity and trading activity, Journal of Finance 56, 501-530.

Chordia, T., R. Roll, and A. Subrahmanyam, 2007, Liquidity and the law of one price: The case of the futures-cash basis, Journal of Finance 62, 2201-2234.

Chordia, T., A. Subrahmanyam, and Q. Tong, 2014, Have Capital Market Anomalies Attenuated in the Recent Era of High Liquidity and Trading Activity?, Journal of Accounting and Economics 58, 41-58.

Comerton-Forde, C., T. Hendershott, C. Jones, P. Moulton, and M. Seasholes, 2010, Time variation in liquidity: The role of market-maker inventories and revenues, Journal of Finance 55, 295-331.

Duarte, J., and L. Young, 2009, Why is PIN priced?, Journal of Financial Economics 91, 119-138.

Easley, D., S. Hvidkjaer, and M. O'Hara, 2002, Is information risk a determinant of asset returns?, Journal of Finance 57, 2185-2221.

Easley, D., M. Lopez de Prado, and M. O'Hara, 2012, Bulk classification of trading activity, Working paper, Cornell University.

Easley, D, and M. O'Hara, 2004, Information and the cost of capital, Journal of Finance 59, 1553-1583.

Fama, E., and K. French, 1992, The cross-section of expected stock returns, Journal of Finance 47, 427-466.

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

Fama, E., and J. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, Journal of Political Economy 81, 607-636.

Foster, D., and S. Viswanathan, 1993, Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models, Journal of Finance 48, 187-211.

Garleanu, N., and L. Pedersen, 2004, Adverse selection and the required return, Review of Financial Studies 17, 643-665.

Glosten, L., and L. Harris, 1988, Estimating the components of the bid-ask spread, Journal of Financial Economics 21, 123-142.

Grossman, S., and J. Stiglitz, 1980, On the impossibility of informationally efficient markets, American Economic Review 70, 393-408.

Han, Y., T. Hu, and D. Lesmond, 2012, Liquidity and the pricing of cross-sectional idiosyncratic volatility around the world, Working paper, Tulane University.

Harvey, C., Y. Liu, and H. Zhu, 2013, …and the cross-section of expected stock returns, Working paper, Duke University.

Hasbrouck, J., 1999, The dynamics of discrete bid and ask quotes, Journal of Finance 54, 2109-2142.

Hasbrouck, J., 2005, Trading costs and returns for US equities: The evidence from daily data, Working paper, New York University.

Hasbrouck, J., 2009, Trading costs and returns for US equities: Estimating effective costs from daily data, Journal of Finance 64, 1445-1477.

Hendershott, T. and A. J. Menkveld, 2014, Price pressures, Journal of Financial Economics 114, 405-423.

Hendershott, T. and M. S. Seasholes, 2007, Market maker inventories and stock prices, American Economic Review 97, 210-214.

Holden, C. and S. Jacobsen, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, Journal of Finance 69, 1747-1785.

Huang, R. and H. Stoll, 1997, The components of the bid-ask spread: A general approach, Review of Financial Studies 10, 995-1034.

Hughes, J., J. Liu, and J. Liu, 2007, Information asymmetry, diversification, and cost of capital, Accounting Review 82, 705-729.

Hwang, L., W. Lee, S. Lim, and K. Park, 2013, Does information risk affect the implied cost of equity capital? An analysis of PIN and adjusted PIN, Journal of Accounting and Economics 55, 148-167.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, Journal of Finance 48, 65-92.

Karolyi, G. A., K. Lee, and M. A. van Dijk, 2012, Understanding commonality in liquidity around the world, Journal of Financial Economics 105, 82–112.

Kelly, B., and A. Ljungqvist, 2012, Testing asymmetric-information asset pricing models, Review of Financial Studies 25, 1366-1413.

Kyle, A., 1985, Continuous auctions and insider trading, Econometrica 53, 1315-1335.

Lai, S., L. Ng, and B. Zhang, 2014, "Does PIN affect equity prices around the world? Journal of Financial Economics, forthcoming.

Lambert, R., C. Leuz, and R. Verrecchia, 2012, Information asymmetry, information precision, and the cost of capital, Review of Finance 16, 1-29.

Lambert, R., and R. Verrecchia, 2014, Information, illiquidity, and cost of capital. Contemporary Accounting Research, forthcoming.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733-747.

Lee, C., and B. Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, Journal of Financial Markets 3, 83-111.

Lee, C., and B. Swaminathan. 2000. Price momentum and trading volume. Journal of Finance 60: 2017-2069.

Levi, S., and X.-J. Zhang, 2014, Do temporary increases in information asymmetry affect the cost of equity?, Management Science, Forthcoming.

Liu, H., and Y. Wang, 2012, Over-the-counter markets: Market making with asymmetric information, inventory risk, and imperfect competition, Working paper, Washington University.

Livnat, J., and R. Mendenhall, 2006, Comparing the post-earnings announcement drift for surprises calculated from analysts and time series forecasts, Journal of Accounting Research 44, 177-205.

Lou, X., and T. Shu, 2014, Why is the Amihud (2002) illiquidity measure priced?, Working paper, University of Georgia.

Mohanram, P., and S. Rajgopal, 2009, Is PIN priced risk?, Journal of Accounting and Economics 47, 226-243.

Nagel, S., 2012, Evaporating liquidity. Review of Financial Studies 25, 2005-2039.

Newey, W., and K. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703-708.

Newey, W., and K. West, 1994, Automatic lag selection in covariance matrix estimation, Review of Economic Studies 61, 631-653.

Odders-White, E., 2000, On the occurrence and consequences of inaccurate trade classification, Journal of Financial Markets 3, 205-332.

O'Hara, M., C. Yao, and M. Ye, 2011, What's not there: The odd-lot bias in TAQ data, Working paper, Cornell University.

Pástor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, Journal of Political Economy 113, 642-685.

Roll, R., 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, Journal of Finance 39, 1127-1139.

Sadka, R., 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, Journal of Financial Economics 80, 309-349.

Stoll, H., 2000, Friction, Journal of Finance 55, 1479-1514.

Stoll, H., 2014, High speed equities trading: 1993-2012, Asia-Pacific Journal of Financial Studies 43, 767-797.

U.S. Securities and Exchange Commission, 2012, Report to Congress on Decimalization, Washington, DC.

Yan, Y. and S. Zhang, 2012, An improved estimation method and empirical properties of the probability of informed trading, Journal of Banking and Finance 36, 454-467.

Table 1
Summary of the Intra-Daily Order Flow Data

This table summarizes key attributes of the transaction-level order flow data used in our study sample. We obtain intraday transaction data from the Institute for the Study of Security Markets (ISSM) for the 1983-1992 period and the NYSE Trades and Automated Quotations (TAQ) for the 1993-2010 period. Trades and quotes are matched based on the Lee and Ready (1991) method up to December 2006 and based on the Holden and Jacobsen (2014) algorithm from January 2007 to December 2010. We limit our study sample to stocks listed on the NYSE or AMEX because of different trading protocols and data availability for NASDAQ stocks. We use only those stocks with at least 50 trades per month. Trades executed at the quote midpoint are not counted. We exclude the trades that are out of sequence, recorded before the open or after the close, or involved in errors or corrections. Quotes before the open or after the close are also excluded.

Total Number of Trades (and Matched Quotes) Used over the Sample Period	17,067,629,571
Minimum Number of Monthly Trades (and Matched Quotes) per Stock	50
Maximum Number of Monthly Trades (and Matched Quotes) per Stock	14,046,042
Average Number of Monthly Trades (and Matched Quotes) per Stock	26,980
Pooled Number of Firms Used over the 336 Months (Firm-month Observations)	632,614

Table 2
Descriptive Statistics of the Non-Information Cost Measures and Other Variables

Panel A reports the descriptive statistics for the four (raw) non-information cost measures. They are defined as follows. $\bar{\varphi}_0^{GH} \equiv$ the Glosten and Harris (1988) non-information component; $\bar{\varphi}_0^{FV} \equiv$ the Foster and Viswanathan (1993) non-information component; $\bar{\varphi}_0^{S} \equiv$ the Sadka (2006) transitory fixed cost; and $\bar{\lambda}_0^{S} \equiv$ the Sadka (2006) transitory variable cost. Since the cost components are small, $\bar{\varphi}_0^{i}$ is multiplied by 10^2 and $\bar{\lambda}_0^{i}$ is multiplied by 10^6 before scaling or Winsorizing. We first calculate the cross-sectional mean, median, standard deviation (*STD*), coefficient of variation (*CV*), skewness, and kurtosis for each variable in each month and then obtain the time-series average of these values. The same statistics for the price-scaled values (i.e., $\frac{\bar{\varphi}_0^{GH}}{P}, \frac{\bar{\varphi}_0^{FV}}{P}, \frac{\bar{\varphi}_0^{S}}{P}$, and $\frac{\bar{\lambda}_0^{S}}{P}$) of $\bar{\varphi}_0^{GH}, \bar{\varphi}_0^{FV}, \bar{\varphi}_0^{S}$, and $\bar{\lambda}_0^{S}$ are calculated using the previous month-end stock price (*P*) and reported in Panel B. Panel C reports descriptive statistics for the Winsorized values (at the 0.5th and 99.5th percentiles) ($\bar{\varphi}^{GH}, \bar{\varphi}^{FV}, \bar{\varphi}^{S}$, and $\bar{\lambda}^{S}$) of $\bar{\varphi}_0^{GH}, \bar{\varphi}_0^{FV}, \bar{\varphi}_0^{S}$, and $\bar{\lambda}_0^{S}$ and Panel D does the same for the Winsorized values ($\frac{\bar{\varphi}^{GH}}{P}, \frac{\bar{\varphi}^{FV}}{P}, \frac{\bar{\varphi}^{S}}{P}$, and $\frac{\bar{\lambda}^{S}}{P}$) of $\frac{\bar{\varphi}_0^{GH}}{P}, \frac{\bar{\varphi}_0^{FV}}{P}, \frac{\bar{\varphi}_0^{S}}{P}$, and $\frac{\bar{\lambda}_0^{S}}{P}$. Panel E reports the descriptive statistics for the following firm attributes. *SIZE* ≡ the natural logarithm of the previous month-end market value of equity (*MV*); *BM* ≡ the book-to-market ratio (*BM* = *BV*/*MV*), where *BV* and *MV* are the book and market values of equity in million dollars; *BTM* ≡ the Winsorized value (at the 0.5th and 99.5th percentiles) of *BM;* and *PAR1*, *PAR2*, *PAR3*, and *PAR4* ≡ the compounded holding period returns over the most recent three months (from month *t*-1 to month *t*-3), from month *t*-4 to month *t*-6, from month *t*-7 to month *t*-9, and from month *t*-10 to month *t*-12, respectively. For each of the momentum variables to exist, a stock must have all three returns over the corresponding three-month period. We calculate the book-to-market ratio using data from the CRSP and CRSP/Compustat Merged (CCM) quarterly files, and calculate other variables using data from the CRSP monthly file. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. The average number of component stocks used each month is 1,888.

Panel A: Descriptive Statistics for the Non-Information Cost Measures

Non-Information Cost Measure	Mean	Median	STD	CV	Skewness	Kurtosis
$\bar{\varphi}_0^{GH}$	4.318	3.638	21.527	243.89	17.72	662.32
$\bar{\varphi}_0^{FV}$	4.268	3.631	19.884	245.04	17.52	665.42
$\bar{\varphi}_0^{S}$	3.214	3.026	13.577	449.20	3.35	689.17
$\bar{\lambda}_0^{S}$	0.233	-0.016	11.207	2132.38	12.96	681.09

Panel B: Descriptive Statistics for the Price-Scaled Non-Information Cost Measures

Non-Information Cost Measure	Mean	Median	STD	CV	Skewness	Kurtosis
$\bar{\varphi}_0^{GH}/P$	0.422	0.179	5.016	425.65	9.19	364.02
$\bar{\varphi}_0^{FV}/P$	0.421	0.179	4.989	208.28	9.09	362.88
$\bar{\varphi}_0^{S}/P$	0.234	0.141	3.313	-1163.59	2.04	378.51
$\bar{\lambda}_0^{S}/P$	0.679	-0.001	23.406	-1120.51	8.65	727.12

(Table 2: continued)

Non-Information Cost Measure	Mean	Median	STD	CV	Skewness	Kurtosis
Panel C: Descriptive Statistics for the Winsorized Non-Information Cost Measures						
$\bar{\varphi}^{GH}$	3.829	3.638	1.392	62.68	3.50	27.40
$\bar{\varphi}^{FV}$	3.819	3.631	1.388	62.55	3.50	27.57
$\bar{\varphi}^{S}$	3.042	3.026	1.213	66.81	2.31	23.04
$\bar{\lambda}^{S}$	-0.069	-0.016	0.669	-1111.46	0.29	26.99
Panel D: Descriptive Statistics for the Price-Scaled and Winsorized Non-Information Cost Measures						
$\bar{\varphi}^{GH}/P$	0.312	0.179	0.474	160.60	4.52	26.76
$\bar{\varphi}^{FV}/P$	0.311	0.179	0.474	161.05	4.54	26.94
$\bar{\varphi}^{S}/P$	0.257	0.141	0.396	173.82	4.62	27.66
$\bar{\lambda}^{S}/P$	-0.004	-0.001	0.373	-1747.47	-1.56	52.01

Panel E: Descriptive Statistics for Other Variables

Variables	Mean	Median	STD	CV	Skewness	Kurtosis
MV	3,386.90	644.64	11,311.77	319.67	9.88	145.84
SIZE	6.30	6.31	1.86	29.63	0.01	-0.22
BM	0.69	0.59	0.62	87.23	6.57	115.39
BTM	0.69	0.59	0.53	74.87	2.74	14.14
PAR1	0.013	0.020	0.200	-1441.35	-0.39	8.02
PAR2	0.014	0.021	0.195	26.07	-0.30	6.98
PAR3	0.016	0.022	0.192	-138.69	-0.23	6.61
PAR4	0.017	0.022	0.190	-506.67	-0.15	6.58

Table 3
Correlations between the Variables

This table reports the correlations (the time-series average of monthly cross-sectional correlation coefficients) between the key variables of interest in our study. $\frac{\bar{\varphi}^{GH}}{P}$, $\frac{\bar{\varphi}^{FV}}{P}$, $\frac{\bar{\varphi}^{S}}{P}$, and $\frac{\bar{\lambda}^{S}}{P}$ are the Winsorized (at the 0.5th and 99.5th percentiles) and price-scaled values (by the previous month-end price, P) of the following variables. $\bar{\varphi}_0^{GH} \equiv$ the Glosten and Harris (1988) non-information component; $\bar{\varphi}_0^{FV} \equiv$ the Foster and Viswanathan (1993) non-information component; $\bar{\varphi}_0^{S} \equiv$ the Sadka (2006) transitory fixed cost; and $\bar{\lambda}_0^{S} \equiv$ the Sadka (2006) transitory variable cost. Since the cost components are small, $\bar{\varphi}_0^{i}$ is multiplied by 10^2 and $\bar{\lambda}_0^{i}$ is multiplied by 10^6 before Winsorizing or scaling. Other variables are defined as follows. *SIZE* ≡ the natural logarithm of the previous month-end market value of equity (*MV*) in million dollars; *BTM* ≡ the Winsorized value (at the 0.5th and 99.5th percentiles) of the book-to-market ratio (*BM*); and *PAR1*, *PAR2*, *PAR3*, and *PAR4* ≡ the compounded holding period returns over the most recent three months (from month *t*-1 to month *t*-3), from month *t*-4 to month *t*-6, from month *t*-7 to month *t*-9, and from month *t*-10 to month *t*-12, respectively. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. The average number of component stocks used each month is 1,888.

	$\bar{\varphi}^{GH}/P$	$\bar{\varphi}^{FV}/P$	$\bar{\varphi}^{S}/P$	$\bar{\lambda}^{S}/P$	SIZE	BTM	PAR1	PAR2	PAR3	PAR4
$\bar{\varphi}^{GH}/P$	1									
$\bar{\varphi}^{FV}/P$	0.999	1								
$\bar{\varphi}^{S}/P$	0.956	0.957	1							
$\bar{\lambda}^{S}/P$	-0.191	-0.192	-0.305	1						
SIZE	-0.564	-0.562	-0.506	0.108	1					
BTM	0.235	0.235	0.223	-0.053	-0.300	1				
PAR1	-0.201	-0.201	-0.196	0.045	0.119	0.012	1			
PAR2	-0.189	-0.189	-0.186	0.048	0.118	-0.017	0.032	1		
PAR3	-0.177	-0.178	-0.174	0.043	0.114	-0.145	0.043	0.025	1	
PAR4	-0.163	-0.163	-0.160	0.038	0.107	-0.164	0.035	0.038	0.020	1

Table 4
Results of Monthly Cross-Sectional Regressions with the Non-Information Cost Measures

Panels A and B report the Fama-MacBeth (1973) regression results using the two non-information (transitory) cost measures: one ($\frac{\bar{\varphi}^{GH}}{P}$) based on Glosten and Harris (1998) and the other ($\frac{\bar{\varphi}^{FV}}{P}$) based on Foster and Viswanathan (1993). Panel C reports the regression results using the components further decomposed based on Sadka (2006): the transitory fixed cost ($\frac{\bar{\varphi}^{S}}{P}$) and the transitory variable cost ($\frac{\bar{\lambda}^{S}}{P}$). The return to be used as the dependent variable is five-factor (F5) adjusted with regard to the Fama-French three factors (MKT_t, SMB_t, and HML_t), Carhart's (1997) momentum factor (UMD_t), and Pastor-Stambaugh's (2003) liquidity factor (LIQ_t) in two ways. In the first method, we calculate the F5-adjusted return (R_{jt}^{e1}) in each month via the formula, $R_{jt}^{e1} = (\tilde{R}_{jt} - R_{Ft}) - (\hat{\beta}_{j1}^{*}MKT_t + \hat{\beta}_{j2}^{*}SMB_t + \hat{\beta}_{j3}^{*}HML_t + \hat{\beta}_{j4}^{*}UMD_t + \hat{\beta}_{j5}^{*}LIQ_t)$, after estimating the factor loadings for each stock with the entire sample range of the data (from January 1983 to December 2010). In the second method, we obtain the rolling estimates of the factor loadings, β_{jk} (k =1 to 5), in each month using the time series of the past 60 months. Given the current month's data ($\tilde{R}_{jt} - R_{Ft}$, MKT_t, SMB_t, HML_t, UMD_t, and LIQ_t) and the estimated factor loadings ($\hat{\beta}_{jk}^{**}$), we calculate the second F5-adjusted return, R_{jt}^{e2}, in each month via the formula, $R_{jt}^{e2} = (\tilde{R}_{jt} - R_{Ft}) - (\hat{\beta}_{j1}^{**}MKT_t + \hat{\beta}_{j2}^{**}SMB_t + \hat{\beta}_{j3}^{**}HML_t + \hat{\beta}_{j4}^{**}UMD_t + \hat{\beta}_{j5}^{**}LIQ_t)$. Using the F5-adjusted returns, we then estimate the Fama-MacBeth (1973) cross-sectional regression, $R_{jt}^{eh} = c_{0t-1} + \gamma\Lambda_{jt-1}^{i} + \sum_{n=1}^{N} c_{nt}X_{njt-1} + \tilde{e}_{jt}^{\prime}$, where h = 1 or 2, Λ_{jt-1}^{i} is the non-information component of trading costs, and X_{njt-1} denotes firm characteristics (n = 1, …, N) for security j in month *t-1*. For comparison purposes, we also report the result from using the F5-*unadjusted* excess returns (R^e). The coefficient vector $\boldsymbol{c}_t = [c_{0t}\ \gamma_t\ c_{1t}\ c_{2t} \dots c_{Nt}]'$ is estimated each month with ordinary least-squares (OLS) regressions. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates. Since the cost components are small, $\bar{\varphi}^{i}$ is multiplied by 10^2 and $\bar{\lambda}^{S}$ is multiplied by 10^6 before Winsorizing or scaling (by the previous month-end price, P). We include the following firm attributes (X_{njt-1}) in the regression. *SIZE* ≡ the natural logarithm of the (previous month-end) market value of equity (MV) in million dollars; *BTM* ≡ the book-to-market ratio Winsorized at the 0.5th and 99.5th percentiles; and *PAR1*, *PAR2*, *PAR3*, and *PAR4* ≡ the compounded holding period returns of a stock over the most recent three months (from month *t*-1 to month *t*-3), from month *t*-4 to month *t*-6, from month *t*-7 to month *t*-9, and from month *t*-10 to month *t*-12, respectively. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. In addition to the average coefficients and *t*-statistics (italic), we also report the mean value of the adjusted R^2 values from the individual regressions (*Avg R-sqr*) and the mean number of stocks (*Avg Obs*) used in the regressions. The coefficients are multiplied by 100. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table 4: continued)

Panel A: Non-Information Cost Measure Based on Glosten and Harris (1988)

Expla. Variables	R^e		R^{e1}		R^{e2}	
Intercept	-0.398		-0.909	**	-0.603	**
	-1.09		*-4.86*		*-2.86*	
$\frac{\bar{\varphi}^{GH}}{P}$	2.755	**	2.809	**	2.977	**
	9.46		*9.91*		*9.46*	
SIZE	0.034		0.032		0.006	
	0.98		*1.49*		*0.25*	
BTM	0.248	*	0.176	*	0.090	
	2.44		*2.27*		*1.02*	
PAR1	-0.066		-0.263		-0.910	*
	-0.18		*-0.92*		*-2.44*	
PAR2	1.193	**	1.352	**	1.457	**
	3.82		*5.67*		*4.71*	
PAR3	1.587	**	1.540	**	1.237	**
	5.50		*6.59*		*3.64*	
PAR4	1.506	**	1.378	**	1.205	**
	5.92		*6.69*		*3.98*	
Avg R-sqr	0.061		0.041		0.049	
Avg Obs	1,844		1,844		1,686	

Panel B: Non-Information Cost Measure Based on Foster and Viswanathan (1993)

Expla. Variables	R^e		R^{e1}		R^{e2}	
Intercept	-0.359		-0.870	**	-0.568	**
	-0.99		*-4.67*		*-2.70*	
$\frac{\bar{\varphi}^{FV}}{P}$	2.722	**	2.778	**	2.955	**
	9.36		*9.81*		*9.38*	
SIZE	0.029		0.027		0.001	
	0.85		*1.28*		*0.07*	
BTM	0.248	*	0.177	*	0.091	
	2.45		*2.27*		*1.03*	
PAR1	-0.075		-0.271		-0.916	*
	-0.20		*-0.94*		*-2.45*	
PAR2	1.185	**	1.344	**	1.450	**
	3.79		*5.64*		*4.69*	
PAR3	1.579	**	1.533	**	1.231	**
	5.47		*6.56*		*3.62*	
PAR4	1.498	**	1.371	**	1.199	**
	5.89		*6.65*		*3.96*	
Avg R-sqr	0.061		0.041		0.049	
Avg Obs	1,844		1,844		1,686	

(Table 4: continued)

Panel C: Non-Information Cost Measures Based on Sadka (2006)

Explanatory	R^e			R^{e1}			R^{e2}		
Variables	1	2	3	4	5	6	7	8	9
Intercept	0.391	1.680 **	0.348	-0.129	1.139 **	-0.172	0.161	1.502 **	0.111
	1.05	*4.18*	*0.93*	*-0.72*	*5.41*	*-0.95*	*0.79*	*6.28*	*0.54*
$\frac{\bar{\varphi}^S}{P}$	2.101 **		2.225 **	2.195 **		2.325 **	2.350 **		2.529 **
	6.79		*6.86*	*7.33*		*7.42*	*7.01*		*7.14*
$\frac{\bar{\lambda}^S}{P}$		-0.714 *	0.203		-0.744 **	0.241		-0.421	0.559 *
		-2.49	*0.62*		*-2.58*	*0.75*		*-1.64*	*2.24*
SIZE	-0.057	-0.206 **	-0.052	-0.058 **	-0.204 **	-0.053 *	-0.082 **	-0.236 **	-0.077 **
	-1.62	*-5.00*	*-1.47*	*-2.77*	*-7.79*	*-2.52*	*-3.56*	*-8.18*	*-3.29*
BTM	0.272 **	0.327 **	0.266 **	0.200 **	0.258 **	0.196 *	0.115	0.173	0.108
	2.69	*3.21*	*2.63*	*2.58*	*3.32*	*2.53*	*1.30*	*1.93*	*1.23*
PAR1	-0.273	-0.642	-0.257	-0.463	-0.834 **	-0.446	-1.101 **	-1.497 **	-1.083 **
	-0.73	*-1.68*	*-0.69*	*-1.60*	*-2.81*	*-1.54*	*-2.93*	*-3.90*	*-2.88*
PAR2	1.001 **	0.657 *	1.026 **	1.167 **	0.823 **	1.190 **	1.273 **	0.935 **	1.300 **
	3.20	*2.07*	*3.27*	*4.93*	*3.45*	*5.00*	*4.14*	*3.07*	*4.21*
PAR3	1.427 **	0.980 **	1.421 **	1.389 **	0.955 **	1.385 **	1.102 **	0.664	1.092 **
	4.93	*3.29*	*4.93*	*5.94*	*3.94*	*5.96*	*3.24*	*1.89*	*3.22*
PAR4	1.369 **	1.125 **	1.377 **	1.241 **	1.001 **	1.251 **	1.078 **	0.816 **	1.081 **
	5.38	*4.34*	*5.41*	*6.03*	*4.81*	*6.10*	*3.57*	*2.70*	*3.59*
Avg R-sqr	0.059	0.052	0.061	0.039	0.032	0.041	0.047	0.041	0.049
Avg Obs	1,844	1,843	1,843	1,844	1,843	1,843	1,686	1,685	1,685

Table 5
Results with the Sample from the TAQ Period Only and Excluding the High-Frequency-Trading Years

Panels A and B report the Fama-MacBeth (1973) regression results with the sample from the TAQ period (1993:01-2010:12) only, and Panels C and D do the same with the sample excluding the high-frequency-trading years (2007:01-2010:12). Panels A and C use the measure ($\frac{\bar{\varphi}^{GH}}{P}$), which is based on Glosten and Harris (1998), while Panels B and D use the measure ($\frac{\bar{\varphi}^{FV}}{P}$) based on Foster and Viswanathan (1993). The return to be used as the dependent variable is F5-adjusted with regard to the Fama-French three factors (MKT_t, SMB_t, and HML_t), Carhart's (1997) momentum factor (UMD_t), and Pastor-Stambaugh's (2003) liquidity factor (LIQ_t) in two ways as described in the previous tables. Using the F5-adjusted returns (R_{jt}^{e1} and R_{jt}^{e2}), we estimate the Fama-MacBeth (1973) cross-sectional regression, $R_{jt}^{eh} = c_{0t-1} + \gamma\Lambda_{jt-1}^{i} + \sum_{n=1}^{N} c_{nt}X_{njt-1} + \tilde{e}_{jt}'$, where h = 1 or 2, Λ_{jt-1}^{i} is the non-information component of trading costs, and X_{njt-1} denotes firm characteristics (n = 1, …, N) for security j in month *t-1*. The coefficient vector $\boldsymbol{c_t} = [c_{0t}\ \gamma_t\ c_{1t}\ c_{2t} \dots c_{Nt}]'$ is estimated each month with ordinary least-squares (OLS) regressions. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates. Since the cost components are small, $\bar{\varphi}^{i}$ is multiplied by 10^2 before Winsorizing or scaling (by the previous month-end price, P). The firm attributes (X_{njt-1}) are defined in the previous tables. In addition to the average coefficients and *t*-statistics (italic), we also report the mean value of the adjusted R^2 values from the individual regressions (*Avg R-sqr*) and the mean number of stocks (*Avg Obs*) used in the regressions. The coefficients are multiplied by 100. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

With the Sample from the TAQ Period (1993-2010) Only

Panel A: Glosten and Harris (1988)					Panel B: Foster and Viswanathan (1993)				
Explanatory Variables	R^{e1}		R^{e2}		Explanatory Variables	R^{e1}		R^{e2}	
Intercept	-1.394	**	-1.228	**	Intercept	-1.343	**	-1.184	**
	-6.02		*-4.59*			*-5.82*		*-4.45*	
$\frac{\bar{\varphi}^{GH}}{P}$	3.685	**	3.963	**	$\frac{\bar{\varphi}^{FV}}{P}$	3.641	**	3.934	**
	9.11		*8.73*			*9.01*		*8.65*	
SIZE	0.099	**	0.081	**	SIZE	0.094	**	0.076	**
	3.91		*2.89*			*3.70*		*2.72*	
BTM	0.217	*	0.109		BTM	0.216	*	0.109	
	2.22		*0.96*			*2.22*		*0.97*	
PAR1	-0.163		-0.784		PAR1	-0.172		-0.790	
	-0.45		*-1.59*			*-0.47*		*-1.61*	
PAR2	1.210	**	1.328	**	PAR2	1.200	**	1.318	**
	4.28		*3.62*			*4.25*		*3.60*	
PAR3	1.134	**	0.506		PAR3	1.125	**	0.499	
	3.90		*1.11*			*3.87*		*1.10*	
PAR4	1.385	**	1.513	**	PAR4	1.377	**	1.508	**
	5.92		*3.99*			*5.88*		*3.98*	
Avg R-sqr	0.040		0.051		Avg R-sqr	0.040		0.051	
Avg Obs	2,073		1,896		Avg Obs	2,072		1,896	

(Table 5: continued)

Excluding the High-Frequency-Trading Years (2007-2010)									
Panel C: Glosten and Harris (1988)					**Panel D: Foster and Viswanathan (1993)**				
Explanatory Variables	R^{e1}		R^{e2}		Explanatory Variables	R^{e1}		R^{e2}	
Intercept	-0.586	**	-0.129		Intercept	-0.544	**	-0.092	
	-3.06		*-0.61*			*-2.86*		*-0.43*	
$\frac{\bar{\varphi}^{GH}}{P}$	2.637	**	2.678	**	$\frac{\bar{\varphi}^{FV}}{P}$	2.600	**	2.652	**
	8.87		*8.31*			*8.76*		*8.23*	
SIZE	-0.011		-0.048	*	SIZE	-0.016		-0.052	*
	-0.51		*-2.02*			*-0.73*		*-2.21*	
BTM	0.235	**	0.160		BTM	0.236	**	0.161	
	2.76		*1.70*			*2.78*		*1.72*	
PAR1	-0.111		-0.673		PAR1	-0.120		-0.681	
	-0.39		*-1.81*			*-0.42*		*-1.83*	
PAR2	1.620	**	1.762	**	PAR2	1.611	**	1.754	**
	6.54		*5.45*			*6.51*		*5.42*	
PAR3	1.806	**	1.763	**	PAR3	1.797	**	1.757	**
	7.39		*5.39*			*7.34*		*5.37*	
PAR4	1.472	**	1.165	**	PAR4	1.464	**	1.159	**
	6.53		*3.65*			*6.49*		*3.63*	
Avg R-sqr	0.041		0.048		Avg R-sqr	0.040		0.048	
Avg Obs	1,815		1,648		Avg Obs	1,815		1,648	

Table 6
Controlling for Idiosyncratic Volatility and Earnings Surprise

Panels A and B report the results from the Fama-MacBeth (1973) regressions that control for idiosyncratic volatility (*IVOL*) and earnings surprise (*SUE*): with $\frac{\bar{\varphi}^{GH}}{P}$, in Panel A and with $\frac{\bar{\varphi}^{FV}}{P}$ in Panel B. The return to be used as the dependent variable is F5-adjusted with regard to the Fama-French three factors (MKT_t, SMB_t, and HML_t), Carhart's (1997) momentum factor (UMD_t), and Pastor-Stambaugh's (2003) liquidity factor (LIQ_t) in two ways as described in the previous tables. Using the F5-adjusted returns (R_{jt}^{e1} and R_{jt}^{e2}), we estimate the Fama-MacBeth (1973) cross-sectional regression, $R_{jt}^{eh} = c_{0t-1} + \gamma\Lambda_{jt-1}^{i} + \sum_{n=1}^{N} c_{nt}X_{njt-1} + \tilde{e}_{jt}'$, where h = 1 or 2, Λ_{jt-1}^{i} is the non-information component of trading costs, and X_{njt-1} denotes firm characteristics (n = 1, …, N) for security j in month t-1. The coefficient vector $\boldsymbol{c}_t = [c_{0t}\ \gamma_t\ c_{1t}\ c_{2t} \dots c_{Nt}]'$ is estimated each month with ordinary least-squares (OLS) regressions. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates. Since the cost components are small, $\bar{\varphi}^{i}$ is multiplied by 10^2 before Winsorizing or scaling (by the previous month-end price, P). *IVOL* is the standard deviation of residuals from the time-series regression of the monthly excess return of each stock on the five factors (MKT_t, SMB_t, HML_t, UMD_t, and LIQ_t) using the rolling window of the past 60 months. We measure earnings surprise by $SUE_{it} = \frac{EPS_{it} - EPS_{it-4}}{P_{it}}$, where EPS_{it} is the "street" earnings per share for firm i in quarter t that excludes special items from the Compustat-reported EPS; P_{it} is the stock price at the end of quarter t; and EPS_{it-4} is the EPS at the end of quarter t-4 (adjusted for stock splits and stock dividends). Other firm attributes are defined in the previous tables. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. In addition to the average coefficients and *t*-statistics (italic), we also report the mean value of the adjusted R^2 values from the individual regressions (*Avg R-sqr*) and the mean number of stocks (*Avg Obs*) used in the regressions. The coefficients are multiplied by 100. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table 6: continued)

Panel A: Glosten and Harris (1988)					Panel B: Foster and Viswanathan (1993)				
Explanatory Variables	R^{e1}		R^{e2}		Explanatory Variables	R^{e1}		R^{e2}	
Intercept	0.669	**	0.861	**	Intercept	0.694	**	0.885	**
	3.12		*3.51*			*3.24*		*3.62*	
$\frac{\bar{\varphi}^{GH}}{P}$	3.167	**	3.458	**	$\frac{\bar{\varphi}^{FV}}{P}$	3.135	**	3.427	**
	11.35		*11.20*			*11.25*		*11.11*	
IVOL	*-9.49*	**	*-9.89*	**	IVOL	*-9.39*	**	*-9.79*	**
	-9.32		*-7.94*			*-9.23*		*-7.87*	
SUE	2.064	**	2.134	**	SUE	2.076	**	2.145	**
	4.16		*4.01*			*4.18*		*4.03*	
SIZE	-0.072	**	-0.087	**	SIZE	-0.076	**	-0.091	**
	-3.38		*-3.64*			*-3.57*		*-3.81*	
BTM	0.072		0.001		BTM	0.074		0.003	
	0.95		*0.01*			*0.98*		*0.04*	
PAR1	-0.278		-0.836	*	PAR1	-0.287		-0.846	*
	-0.93		*-2.26*			*-0.96*		*-2.29*	
PAR2	1.358	**	1.556	**	PAR2	1.347	**	1.545	**
	5.43		*5.00*			*5.40*		*4.97*	
PAR3	1.572	**	1.210	**	PAR3	1.565	**	1.203	**
	6.37		*3.56*			*6.33*		*3.54*	
PAR4	1.370	**	1.218	**	PAR4	1.362	**	1.211	**
	6.20		*3.99*			*6.16*		*3.97*	
Avg R-sqr	0.050		0.057		Avg R-sqr	0.050		0.057	
Avg Obs	1,685		1,685		Avg Obs	1,685		1,685	

Table 7
Results under Different Trading Regimes in the NYSE

This table reports the results of the monthly Fama-MacBeth (1973) regressions under different trading regimes for NYSE-listed stocks. The two sub-periods that we consider are the $1/8 era (from January 1983 to May 1997) and the decimal era (from February 2001 to December 2010). The middle period ($ 1/16 regime) from July 1997 to December 2000 is not considered because the interval is too short to calculate the Fama-MacBeth statistics. For this experiment, we exclude AMEX-listed stocks from the study sample. To save space, we report only the results using the Foster and Viswanathan (1993) measure ($\frac{\bar{\varphi}^{FV}}{P}$). Since the cost components are small, $\bar{\varphi}^{FV}$ is multiplied by 10^2 before Winsorizing or scaling (by the previous month-end price, P). Other variables are defined in the previous tables. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates. The coefficients are multiplied by 100. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

	Panel A: $1/8 Era (1983:01-1997:05)				Panel B: Decimalized Era (2001:02-2010:12)			
Expla. Variables	R^{e1}		R^{e2}		R^{e1}		R^{e2}	
Intercept	-0.589	*	-0.179		-1.169	**	-0.873	*
	-2.15		*-0.65*		*-3.54*		*-2.32*	
$\frac{\bar{\varphi}^{FV}}{P}$	0.957	**	0.928	**	6.379	**	6.866	**
	4.56		*4.52*		*5.57*		*5.77*	
SIZE	0.021		-0.007		0.113	**	0.079	*
	0.75		*-0.25*		*3.56*		*2.21*	
BTM	0.008		-0.010		-0.228		-0.399	**
	0.08		*-0.09*		*-1.90*		*-2.75*	
PAR1	-0.608		-0.982	*	-0.146		-0.651	
	-1.45		*-2.14*		*-0.25*		*-0.92*	
PAR2	1.103	**	1.358	**	0.893	*	0.795	
	2.85		*3.02*		*1.96*		*1.46*	
PAR3	1.906	**	1.689	**	0.563		-0.278	
	5.97		*4.35*		*1.30*		*-0.42*	
PAR4	1.414	**	0.706		0.457		0.578	
	4.36		*1.65*		*1.40*		*1.24*	
Avg R-sqr	0.040		0.042		0.040		0.049	
Avg Obs	1,293		1,189		1764		1,675	

Table 8
Proportions of the Non-Information Component and the Adverse-Selection Component

This table reports the proportion of each spread component across the portfolios formed by sorting on firm size (*MV*) as well as for the pooled sample. To form the quintiles, each month the component stocks are split into five portfolios (with the equal number of stocks) after being sorted in ascending order by the previous month market value of equity (MV_{t-1}), and then the cross-sectional mean values of the proportion for each component are computed for each portfolio. The time-series average of the monthly cross-sectional mean values is reported in the table. We estimate the proportion of the non-information component ($ProNIC^i$) and the proportion of the adverse-selection component ($ProASC^i$) for a trade of \$10,000 using the following formula: $ProNIC^i = \frac{\bar{\varphi}^i}{\bar{\varphi}^i+\lambda^i}$ (i = GH or FV), and $ProASC^i = \frac{\lambda^i}{\bar{\varphi}^i+\lambda^i}$ (i = GH or FV), where $\bar{\varphi}^i$ and λ^i are the non-information cost and the adverse-selection cost, respectively, estimated by the Glosten and Harris (GH, 1988) model or by the Foster and Viswanathan (FV, 1993) model. Since the cost components are small, $\bar{\varphi}^i$ is multiplied by 10^2 and λ^i is multiplied by 10^6 before Winsorizing or scaling (by the previous month-end price, P). *MV1* is the smallest size quintile and *MV5* is the largest size quintile. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. The average number of component stocks used in each portfolio in a month is 348.5.

		Glosten and Harris (1988)		Foster and Viswanathan (1993)	
Portfolio/Sample		$ProNIC^{GH}$	$ProASC^{GH}$	$ProNIC^{FV}$	$ProASC^{FV}$
	MV1	0.702	0.298	0.700	0.300
	MV2	0.857	0.143	0.856	0.144
Firm Size	MV3	0.918	0.082	0.917	0.083
	MV4	0.956	0.044	0.956	0.044
	MV5	0.980	0.020	0.979	0.021
Pooled Sample		0.878	0.122	0.877	0.123

Table 9
Relative Importance of the Non-Information Cost vs. the Adverse-Selection Cost

To assess the relative importance of the information and non-information components of trading costs in a multivariate setting, we include in the regression analysis either or both of the two components: adverse-selection cost measure ($\frac{\lambda^{GH}}{P}$ or $\frac{\lambda^{FV}}{P}$) and the non-information cost measure ($\frac{\bar{\varphi}^{GH}}{P}$ or $\frac{\bar{\varphi}^{FV}}{P}$). Panel A reports the regression results when we measure the two components using the Glosten and Harris (1988) model, and Panel B reports the results when we measure the two components using the Foster and Viswanathan (1993) model. $\frac{\lambda^{GH}}{P}$ is the adverse-selection component, which is estimated based on Glosten and Harris (1988) using intradaily dollar order flows available within each month, scaled by the previous month-end stock price (P), and then Winsorized at the 0.5th and 99.5th percentiles. $\frac{\lambda^{FV}}{P}$ is the adverse-selection component, which is estimated based on Foster and Viswanathan (1993) using intra-daily dollar order flows available within each month, scaled by the previous month-end stock price (P), and then Winsorized at the 0.5th and 99.5th percentiles. Since the cost components are small, $\bar{\varphi}^{i}$ is multiplied by 10^2 and λ^{i} is multiplied by 10^6 before Winsorizing or scaling. Other variables are defined in the previous tables. For comparison purposes, regression specifications 1 and 4 in each panel reproduce the regression results for the non-information component that are reported earlier in Table 4. Specifications 2 and 5 in each panel show the results when we include only the adverse-selection (information) component together with the other control variables in the regression model, and specifications 3 and 6 show the results when we include both components (non-information and adverse-selection) together with the same control variables. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table 9: continued)

Panel A: Non-Information Cost vs. Adverse-Selection Cost [Based on Glosten and Harris (1988)]												
	R^{e1}						R^{e2}					
Expla. Variables	1		2		3		4		5		6	
Intercept	-0.909	**	0.778	**	-0.868	**	-0.603	**	1.132	**	-0.542	*
	-4.86		*4.06*		*-4.57*		*-2.86*		*5.14*		*-2.54*	
$\frac{\bar{\varphi}^{GH}}{P}$	2.809	**			2.936	**	2.977	**			3.081	**
	9.91				*10.12*		*9.46*				*9.76*	
$\frac{\lambda^{GH}}{P}$			1.217	**	0.320	*			1.338	**	0.422	*
			6.98		*2.07*				*6.67*		*2.28*	
SIZE	0.032		-0.157	**	0.027		0.006		-0.188	**	-0.001	
	1.49		*-6.96*		*1.25*		*0.25*		*-7.48*		*-0.04*	
BTM	0.176	*	0.236	**	0.156	*	0.090		0.136		0.054	
	2.27		*3.05*		*2.02*		*1.02*		*1.52*		*0.61*	
PAR1	-0.263		-0.698	*	-0.244		-0.910	*	-1.340	**	-0.874	*
	-0.92		*-2.38*		*-0.85*		*-2.44*		*-3.52*		*-2.34*	
PAR2	1.352	**	0.916	**	1.360	**	1.457	**	1.032	**	1.468	**
	5.67		*3.86*		*5.70*		*4.71*		*3.40*		*4.73*	
PAR3	1.540	**	1.076	**	1.543	**	1.237	**	0.779	*	1.241	**
	6.59		*4.50*		*6.63*		*3.64*		*2.22*		*3.64*	
PAR4	1.378	**	1.047	**	1.356	**	1.205	**	0.861	**	1.163	**
	6.69		*5.06*		*6.61*		*3.98*		*2.87*		*3.86*	
Avg R-sqr	0.041		0.035		0.044		0.049		0.044		0.053	
Avg Obs	1,844		1,841		1,841		1,687		1,683		1,683	

(Table 9: continued)

Panel B: Non-Information Cost vs. Adverse-Selection Cost [Based on Foster and Viswanathan (1993)]						
Explanatory	R^{e1}			R^{e2}		
Variables	1	2	3	4	5	6
Intercept	-0.870 **	0.844 **	-0.807 **	-0.568 **	1.199 **	-0.485 *
	-4.67	*4.39*	*-4.26*	*-2.70*	*5.43*	*-2.28*
$\frac{\bar{\varphi}^{FV}}{P}$	2.778 **		2.884 **	2.955 **		3.050 **
	9.81		*10.06*	*9.38*		*9.73*
$\frac{\lambda^{FV}}{P}$		1.163 **	0.336 *		1.265	0.413 *
		6.92	*2.23*		*6.66*	*2.36*
SIZE	0.027	-0.166 **	0.019	0.001	-0.197 **	-0.008
	1.28	*-7.29*	*0.89*	*0.07*	*-7.78*	*-0.36*
BTM	0.177 *	0.236 **	0.152 *	0.091	0.136	0.050
	2.27	*3.04*	*1.97*	*1.03*	*1.52*	*0.57*
PAR1	-0.271	-0.707 *	-0.248	-0.916 *	-1.347 **	-0.874 *
	-0.94	*-2.40*	*-0.86*	*-2.45*	*-3.53*	*-2.34*
PAR2	1.344 **	0.897 **	1.349 **	1.450 **	1.010 **	1.456 **
	5.64	*3.77*	*5.65*	*4.69*	*3.32*	*4.69*
PAR3	1.533 **	1.054 **	1.525 **	1.231 **	0.760 *	1.230 **
	6.56	*4.39*	*6.54*	*3.62*	*2.16*	*3.61*
PAR4	1.371 **	1.027 **	1.339 **	1.199 **	0.843 **	1.152 **
	6.65	*4.96*	*6.52*	*3.96*	*2.81*	*3.83*
Avg R-sqr	0.041	0.035	0.044	0.049	0.044	0.053
Avg Obs	1,844	1,841	1,841	1,686	1,683	1,683

Table 10
Relative Importance of the Non-Information Cost: Standardized Regressions and Trade-Size Adjustments for the Adverse-Selection Component

We estimate standardized coefficients to shed further light on the relative importance of the non-information and adverse-selection components. Standardized coefficients measure changes in stock returns (in number of standard deviations) per one standard deviation increase in the explanatory variables. We estimate monthly standardized coefficients and tabulate the Fama-MacBeth and other statistics in Panel of A and Panel B. To examine whether our results in Table 9 are driven by the multicollinearity problem (i.e., a high correlation between the two components), we replace $\frac{\lambda^{GH}}{P}$ and $\frac{\lambda^{FV}}{P}$ with their trade-size-adjusted equivalents C_q_GH and C_q_FV; where $C_q\_GH = \left(\frac{\lambda^{GH}}{P} * q\right)$, $C_q\_FV = \left(\frac{\lambda^{FV}}{P} * q\right)$, and q is the average trade size within each month (in \$100). We calculate q for each firm by dividing the total dollar trading volume within each month (in \$100) by the total number of trades within each month (i.e., the number of buyer-initiated trades + the number of seller-initiated trades within each month). Panels C and D report the regression results with the trade-size-adjusted adverse-selection component. Since the cost components are small, $\bar{\varphi}^i$ is multiplied by 10^2 and λ^i is multiplied by 10^6 before scaling and Winsorizing. Other variables are defined in the previous tables. For comparison purposes, regression specifications 4a, and 4b in Panel A and Panel B reproduce the regression results for the non-information component that are reported earlier in Table 4. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. The coefficients are all multiplied by 100. *Avg R-sqr* is the average of adjusted R-squared. *Avg Obs* is the monthly average number of companies used in the cross-sectional regressions. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

(Table 10: continued)

With Standardized Regression Coefficients: Dep. Var. = R^{e2}

Panel A: Glosten and Harris (1988)

Expla. Vars.	4a		5a		6a	
Intercept	0.000		0.000		0.000	
	-		-		-	
$\frac{\bar{\varphi}^{GH}}{P}$	7.182	**			7.356	**
	10.62				*11.07*	
$\frac{\lambda^{GH}}{P}$			3.633	**	0.881	*
			7.62		*2.14*	
SIZE	-0.258		-3.120	**	-0.560	
	-0.78		*-8.68*		*-1.69*	
BTM	0.458		0.729	*	0.346	
	1.32		*2.08*		*1.00*	
PAR1	-1.119	*	-1.759	**	-1.116	*
	-2.09		*-3.24*		*-2.09*	
PAR2	2.177	**	1.569	**	2.177	**
	4.96		*3.64*		*4.96*	
PAR3	2.029	**	1.364	**	1.980	**
	4.41		*2.94*		*4.33*	
PAR4	1.759	**	1.302	**	1.711	**
	4.11		*3.08*		*4.02*	
Avg R-sqr	0.049		0.045		0.054	
Avg Obs	1,687		1,679		1,679	

Panel B: Foster and Viswanathan (1993)

Expla. Vars.	4b		5b		6b	
Intercept	0.000		0.000		0.000	
	-		-		-	
$\frac{\bar{\varphi}^{FV}}{P}$	7.078	**			7.273	**
	10.48				*11.00*	
$\frac{\lambda^{FV}}{P}$			3.432	**	0.845	*
			7.43		*2.12*	
SIZE	-0.318		-3.272	**	-0.700	*
	-0.96		*-9.12*		*-2.13*	
BTM	0.467		0.732	*	0.336	
	1.35		*2.09*		*0.97*	
PAR1	-1.130	*	-1.749	**	-1.099	*
	-2.11		*-3.22*		*-2.05*	
PAR2	2.166	**	1.512	**	2.131	**
	4.94		*3.49*		*4.84*	
PAR3	2.020	**	1.330	**	1.953	**
	4.39		*2.85*		*4.26*	
PAR4	1.752	**	1.238	**	1.658	**
	4.09		*2.92*		*3.89*	
Avg R-sqr	0.049		0.045		0.054	
Avg Obs	1,686		1,678		1,678	

(Table 10: continued)

With the Adverse-Selection Cost Adjusted for Trade Size: Dep. Var. = R^{e2}													
Panel C: Glosten and Harris (1988)							Panel D: Foster and Viswanathan (1993)						
Expla. Vars.	4c		5c		6c		Expla. Vars.	4d		5d		6d	
Intercept	-0.603	**	0.055		-1.470	**	Intercept	-0.568	**	0.193		-1.369	**
	-2.86		*0.23*		*-6.51*			*-2.70*		*0.82*		*-6.09*	
$\frac{\bar{\varphi}^{GH}}{P}$	2.977	**			2.358	**	$\frac{\bar{\varphi}^{FV}}{P}$	2.955	**			2.398	**
	9.46				*8.17*			*9.38*				*8.26*	
C_q_GH			0.116	**	0.088	**	C_q_FV			0.107	**	0.083	**
			8.34		*7.19*					*9.08*		*7.86*	
SIZE	0.006		-0.063	*	0.111	**	SIZE	0.001		-0.080	**	0.099	**
	0.25		*-2.32*		*4.61*			*0.07*		*-2.94*		*4.11*	
BTM	0.090		0.181	*	0.090		BTM	0.091		0.178	*	0.086	
	1.02		*2.03*		*1.03*			*1.03*		*2.00*		*0.98*	
PAR1	-0.910	*	-1.447	**	-0.932	*	PAR1	-0.916	*	-1.446	**	-0.926	*
	-2.44		*-3.77*		*-2.49*			*-2.45*		*-3.76*		*-2.47*	
PAR2	1.457	**	0.997	**	1.452	**	PAR2	1.450	**	0.985	**	1.446	**
	4.71		*3.30*		*4.70*			*4.69*		*3.25*		*4.67*	
PAR3	1.237	**	0.720	*	1.226	**	PAR3	1.231	**	0.711	*	1.221	**
	3.64		*2.02*		*3.58*			*3.62*		*2.00*		*3.56*	
PAR4	1.205	**	0.875	**	1.187	**	PAR4	1.199	**	0.862	**	1.180	**
	3.98		*2.93*		*3.97*			*3.96*		*2.89*		*3.95*	
Avg R-sqr	0.049		0.046		0.054		Avg R-sqr	0.049		0.045		0.054	
Avg Obs	1,687		1,687		1,687		Avg Obs	1,686		1,687		1,687	

Table 11
Comparison of the Average *t*-Values and Statistical Significance for the Estimates

This table reports the average *t*-values and the proportions of positive and significant cost elements, which are estimated from the time-series regressions as specified in Eqs. (5) and (7). The cost measures are estimated for each firm each month using intradaily order flows classified based on the Lee and Ready (1991) method up to December 2006 and based on the Holden and Jacobsen (2014) algorithm from January 2007 to December 2010. To survive in the sample, stocks should have at least 50 trades per month (on average 2.5 trades per day). Each cost measure reported in the table is defined as follows. $\bar{\varphi}_0^{GH}$: the non-information cost measure, estimated based on Glosten and Harris (1988) (multiplied by 10^2); $\bar{\varphi}_0^{FV}$: the non-information cost measure, estimated based on Foster and Viswanathan (1993) (multiplied by 10^2); λ_0^{GH}: the adverse-selection cost measure, estimated based on Glosten and Harris (1988) (multiplied by 10^6); and λ_0^{FV}: the adverse-selection cost measure, estimated based on Foster and Viswanathan (1993) (multiplied by 10^6). The sample period is the past 336 months (28 years: 1983:01-2010:12).

Estimation Method	**Non-Information Cost Measures**			**Information-Asymmetry Cost Measures**		
	Notation	Mean *t*-Values	% of Positive & Significant Estimates (5%)	Notation	Mean *t*-Values	% of Positive & Significant Estimates (5%)
Glosten and Harris (1988)	$\bar{\varphi}_0^{GH}$	53.85	99.22%	λ_0^{GH}	8.53	82.44%
Foster and Viswanathan (1993)	$\bar{\varphi}_0^{FV}$	53.53	99.15%	λ_0^{FV}	8.39	81.39%

Table 12
Additional Tests Using a Component of *PIN* or of the Amihud Measure

Panels A and B report the standardized [Fama-MacBeth (1973)] regression results with a component of *PIN* (*PIN_B*), and panels C and D report the results with a component of the Amihud measure (A^-). The table uses the two non-information cost measures: one ($\frac{\bar{\varphi}^{GH}}{P}$) based on Glosten and Harris (1998) and the other ($\frac{\bar{\varphi}^{FV}}{P}$) based on Foster and Viswanathan (1993). The return to be used as the dependent variable is F5-adjusted with regard to the five factors (MKT_t, SMB_t, HML_t, UMD_t, and LIQ_t) in two ways as described in the previous tables. Using the F5-adjusted returns (R_{jt}^{e1} and R_{jt}^{e2}), we estimate the Fama-MacBeth (1973) cross-sectional regressions. The coefficient vector is estimated each month with ordinary least-squares (OLS) regressions. The standard Fama-MacBeth (1973) estimator is the time-series average of the monthly coefficients, and the standard error of this estimator is taken from the time series of monthly coefficient estimates. Since the cost components are small, $\bar{\varphi}^i$ is multiplied by 10^2 before Winsorizing or scaling (by the previous month-end price, P). *PIN_B* (the probability of informed selling) is a component of *PIN*, estimated quarterly as in Brennan et al. (2015a) (the measure is lagged by a quarter before converting to a monthly series). A^- (the half-Amihud measure for down days) is a component of the turnover-version Amihud measure, computed as in Brennan et al. (2013). Other variables are defined in Table 4. The sample period is from January 1983 to December 2010 (28 years) for NYSE/AMEX stocks. We report the mean coefficient, Fama-MacBeth (1973) *t*-statistics (italic), the mean of the adjusted R^2 values (*Avg R-sqr*), and the mean number of stocks (*Avg Obs*) used in the regressions. The coefficients are multiplied by 100. Coefficients significantly different from zero at the significance levels of 1% and 5% are indicated by ** and *, respectively.

With a Component of PIN

Panel A: Glosten and Harris (1988)					Panel B: Foster and Viswanathan (1993)				
Explanatory Variables	R^{e1}		R^{e2}		Explanatory Variables	R^{e1}		R^{e2}	
Intercept	0.000		0.000		Intercept	0.000		0.000	
	-		-			-		-	
$\frac{\bar{\varphi}^{GH}}{P}$	7.697	**	7.195	**	$\frac{\bar{\varphi}^{FV}}{P}$	7.573	**	7.087	**
	11.36		*10.67*			*11.21*		*10.53*	
PIN_B	0.912	**	0.920	**	PIN_B	0.912	**	0.918	**
	3.49		*2.86*			*3.49*		*2.86*	
SIZE	1.025	**	0.391		SIZE	0.955	**	0.329	
	2.97		*1.09*			*2.78*		*0.92*	
BTM	0.721	*	0.538		BTM	0.728	*	0.546	
	2.19		*1.49*			*2.21*		*1.51*	
PAR1	-0.257		-0.989		PAR1	-0.270		-1.001	
	-0.51		*-1.76*			*-0.53*		*-1.78*	
PAR2	2.136	**	2.096	**	PAR2	2.120	**	2.087	**
	5.55		*4.73*			*5.51*		*4.71*	
PAR3	2.689	**	2.202	**	PAR3	2.676	**	2.194	**
	6.55		*4.47*			*6.53*		*4.45*	
PAR4	2.305	**	1.809	**	PAR4	2.294	**	1.802	**
	6.70		*4.19*			*6.67*		*4.17*	
Avg R-sqr	0.041		0.050		Avg R-sqr	0.041		0.050	
Avg Obs	1,729		1,588		Avg Obs	1,729		1,588	

(Table 12: continued)

With a Component of the Amihud Measure									
Panel C: Glosten and Harris (1988)					Panel D: Foster and Viswanathan (1993)				
Explanatory Variables	R^{e1}		R^{e2}		Explanatory Variables	R^{e1}		R^{e2}	
Intercept	0.000		0.000		Intercept	0.000		0.000	
	-		-			-		-	
$\frac{\bar{\varphi}^{GH}}{P}$	7.323	**	6.943	**	$\frac{\bar{\varphi}^{FV}}{P}$	7.189	**	6.825	**
	10.74		*10.13*			*10.59*		*9.99*	
A^-	0.626	*	0.686	*	A^-	0.652	*	0.707	*
	2.27		*2.47*			*2.36*		*2.55*	
SIZE	0.409		-0.167		SIZE	0.340		-0.228	
	1.29		*-0.51*			*1.07*		*-0.71*	
BTM	0.758	*	0.461		BTM	0.766	*	0.471	
	2.30		*1.35*			*2.33*		*1.37*	
PAR1	-0.356		-1.102	*	PAR1	-0.369		-1.113	*
	-0.76		*-2.07*			*-0.79*		*-2.09*	
PAR2	2.153	**	2.179	**	PAR2	2.139	**	2.166	**
	5.57		*4.95*			*5.54*		*4.93*	
PAR3	2.492	**	2.002	**	PAR3	2.480	**	1.992	**
	6.61		*4.37*			*6.58*		*4.35*	
PAR4	2.196	**	1.776	**	PAR4	2.185	**	1.769	**
	6.74		*4.14*			*6.71*		*4.12*	
Avg R-sqr	0.042		0.051		Avg R-sqr	0.042		0.051	
Avg Obs	1,840		1,685		Avg Obs	1,840		1,684	

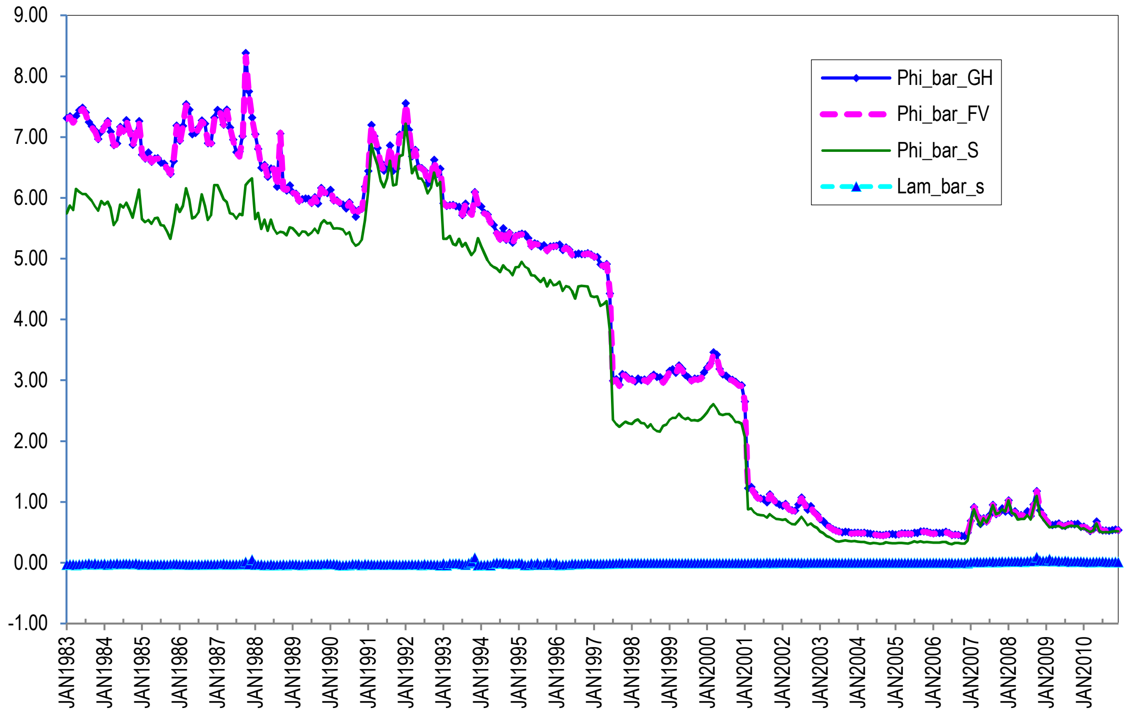

Figure 1. Time-Series Variation in the Non-Information Cost Measures

This figure shows time-series plots of the value-weighted monthly cross-section average for each of the (Winsorized) non-information cost measures for NYSE/AMEX stocks over the past 336 months (28 years: 1983:01-2010:12). The legends in the plot are defined as follws. *Phi_bar_GH* ($\bar{\varphi}^{GH}$): the Glosten and Harris (1988) non-information component; *Phi_bar_FV* ($\bar{\varphi}^{FV}$): the Foster and Viswanathan (1993) non-information component; *Phi_bar_S* ($\bar{\varphi}^{S}$): the Sadka (2006) transitory fixed cost; and *Lam_bar_S* ($\bar{\lambda}^{S}$): the Sadka (2006) transitory variable cost. Since the cost components are small, $\bar{\varphi}^{i}$ is multiplied by 10^2 and $\bar{\lambda}^{i}$ is multiplied by 10^6 before Wisorization.

Appendix: Estimation of the *PIN_B* Measure

Let B_j and S_j be the daily number of buyer- and seller-initiated trades for trading day *j*, respectively. On day *j*, a private information event occurs with probability α, or no information event occurs with probability (1−α). If the event occurs on that day, it contains bad news with probability δ or good news with probability (1−δ). Now orders from uninformed buyers (sellers) arrive randomly at rate ε_b (ε_s) on that day. Also, orders from informed traders arrive randomly at rate μ, but only if the information event occurs on day *j* (i.e., informed traders buy on good news and sell on bad news). Then, the likelihood of observing B_j buys and S_j sells on trading day *j* is given by:

$$L(B_j, S_j|\theta) = \alpha(1-\delta)e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^{B_j}}{B_j!}e^{-\varepsilon_s}\frac{\varepsilon_s^{S_j}}{S_j!} + \alpha\delta e^{-\varepsilon_b}\frac{\varepsilon_b^{B_j}}{B_j!}e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^{S_j}}{S_j!} + (1-\alpha)e^{-\varepsilon_b}\frac{\varepsilon_b^{B_j}}{B_j!}e^{-\varepsilon_s}\frac{\varepsilon_s^{S_j}}{S_j!}, \quad \text{(A1)}$$

where $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ is a vector of the parameters defined above. Assuming that trading days are independent, the joint likelihood of observing a series of daily buys and sells over trading days *j* = 1, 2, …, *J* is the product of the daily likelihoods:

$$L(M|\theta) = \prod_{j=1}^{J} L(\theta|B_j, S_j), \quad \text{(A2)}$$

where M = ((B_1, S_1), …, (B_J, S_J)) is the data set. The parameter vector θ is estimated *quarterly* by maximizing the joint likelihood defined in Eq. (A2) via the Yan and Zhang (2012) method. Using the parameters estimated above, we then calculate the *PIN* measure as follows:

$$PIN = \frac{\alpha\mu}{\alpha\mu+\varepsilon_b+\varepsilon_s}. \quad \text{(A3)}$$

PIN in Eq. (A3) does not distinguish between informed buying on good news and informed selling on bad news. Following Brennan, Huh, and Subrahmanyam (2015a), we thus decompose *PIN* into the probabilities of informed trading based on good news (*PIN_G*) and informed trading based on bad news (*PIN_B*) as follows:

$$PIN = \frac{\alpha\mu}{\alpha\mu+\varepsilon_b+\varepsilon_s} = PIN\_G + PIN\_B,$$

where

$$PIN\_G = \frac{\alpha\mu(1-\delta)}{\alpha\mu+\varepsilon_b+\varepsilon_s}, \quad \text{(A4)}$$

and

$$PIN\_B = \frac{\alpha\mu\delta}{\alpha\mu+\varepsilon_b+\varepsilon_s}. \quad \text{(A5)}$$