The Journal of Financial Research • Vol. XXVIII, No. 3 • Pages 461–485 • Fall 2005

THE IMPACT OF MARKET MAKER CONCENTRATION ON ADVERSE-SELECTION COSTS FOR NASDAQ STOCKS

Bonnie F. Van Ness and Robert A. Van Ness
University of Mississippi

Richard S. Warr
North Carolina State University

Abstract

We examine the impact of market maker concentration on adverse-selection costs for NASDAQ stocks and find that more market makers results in lower costs. Furthermore, this reduction in adverse selection exceeds the overall reduction in spreads that is attributable to market maker competition. We hypothesize that order flow internalization is increasing in market makers and allows for greater information production, and is an explanation for our findings. Our results provide an explanation for the puzzle documented by previous work that finds that adverse-selection costs for NASDAQ tend to be lower than for the New York Stock Exchange, whereas spreads tend to be higher.

JEL Classifications: G14, G18

I. Introduction

In this article we address a puzzling contradiction in the finance literature involving trading costs and level of information transparency for the NASDAQ and New York Stock Exchange (NYSE) markets. Previous research finds that the costs of trading securities are generally higher on dealer markets (such as NASDAQ) than on auction markets (such as the NYSE). For example, Huang and Stoll (1996) find that execution costs (as measured by the quoted, effective, and realized spread) on NASDAQ are twice those on the NYSE. Petersen and Fialkowski (1994) find that execution costs are higher on competing exchanges compared with the NYSE, and Barclay (1997) finds that trading costs decrease after stocks move to the NYSE.

We would like to thank Asli Ascioglu, Kee Chung, Tim McCormick, William T. Moore (the editor), Michael Goldstein, Tom McInish, Andriy Shkilko, Shawn Thomas, Pankaj Jain (the referee), and seminar participants at the University of Memphis and Financial Management Association 2002 annual meetings for many helpful comments and suggestions, and Tim McCormick of NASDAQ who generously provided the market maker volume data. Robert Van Ness acknowledges support from a Hearin Foundation Grant, School of Business Administration, University of Mississippi.

The general consensus of this literature is that trading costs are higher on NASDAQ than on the NYSE, and NASDAQ is a less transparent market.

In contrast to the preceding work, Affleck-Graves, Hegde, and Miller (1994) and Lin, Sanger, and Booth (1998) find that the adverse-selection component of the bid-ask spread is significantly smaller on NASDAQ than on the NYSE. This finding implies that the cost of trading with informed traders is higher on the NYSE than on NASDAQ.

Therefore, the literature presents a contradiction: adverse-selection costs for NASDAQ tend to be lower than for the NYSE, whereas spreads and the other measures of liquidity tend to be higher. The explanation for this contradiction is of importance to investors, companies listing on exchanges, regulators, and the designers of future trading systems. In this article we seek to address this contradiction by examining the effect of a feature of dealer markets that is not present in auction markets, namely, the presence of multiple market makers, on adverse-selection costs. We find that the presence of more market makers results in lower adverse selection as a percentage of the spread for NASDAQ stocks. This finding suggests that it is partly the presence of multiple market makers that results in the lower adverse-selection costs for NASDAQ stocks versus NYSE stocks. The question is: Why does the presence of multiple market makers reduce adverse-selection costs?

There is ample evidence that increased numbers of market makers result in lower spreads. This is consistent with a functioning competitive market. But competition cannot explain why the adverse-selection component of the spread declines with an increase in the number of market makers. A possible explanation may lie in the role that order preferencing plays in the NASDAQ market. Previous work documents that order-preferencing arrangements facilitate information transfer between brokers and dealers, thus reducing the level of asymmetric information faced by the dealer. Such a reduction in asymmetric information would manifest itself in lower adverse-selection costs. Internalization arrangements, where orders are routed through vertically linked broker-dealer firms, provide direct communication between the broker and dealer, allowing the dealer to know the reason behind the trade. Using a simple model, we hypothesize that the prevalence of these internalization arrangements is increasing in the number of market makers.

Our results suggest that greater internalized order flow, which is positively correlated with the number of market makers, reduces adverse-selection costs. Conversely, for stocks with few market makers (and a presumably low proportion of internalized orders), adverse selection is higher and closer to the level of the average NYSE stock in our sample. Our results are economically significant and suggest that although the overall NASDAQ market may not be as transparent as the NYSE, from the market makers' viewpoint, it is.

Our study serves several purposes. First, we add to the literature comparing the NYSE and NASDAQ by examining the determinants of the differences in the level of adverse selection. Second, we show that the adverse-selection components are influenced by similar factors on the two exchanges, but to different degrees, and therefore are not necessarily comparable across exchanges. We believe this incompatibility to be important because several authors use the adverse-selection component as an explanatory variable (e.g., Flannery, Kwan, and Nimalendran 2004). Finally, we contribute to the spread decomposition literature by showing that spread models may perform differently on the various exchanges. Specifically, it appears that the degree of market maker competition (as measured by the number of NASDAQ market makers and the Herfindahl index of market concentration) is an important determinant of adverse selection for NASDAQ stocks. Future researchers using spread decomposition models should attempt to control for this exchange dynamic and control for the concentration of market makers in their studies.

II. Relation Between Market Makers and Adverse Selection

Affleck-Graves, Hegde, and Miller (1994) and Lin, Sanger, and Booth (1998) find that adverse-selection components are significantly lower for NASDAQ stocks than for NYSE stocks. Several explanations exist for why adverse-selection costs are different on dealer versus auction markets. Differences may lie in the structure of the markets, in the types of stocks trading on these markets, or in the participants of the markets. In this section we develop a simple model of the effect that internalization arrangements and market maker concentration have on the level of adverse selection.

Contrary to the NYSE, competition and negotiation allow NASDAQ market makers to know more about the stock being traded and the motivation for the trade. For example, Huang and Stoll (1996) argue that NASDAQ market makers have several significant advantages over the NYSE specialist. First, NASDAQ dealers "know their order flow" through preferencing arrangements and the development of long-term relationships.[1] Second, NASDAQ dealers do not operate independently of one another as negotiation takes place between dealers, and as a result, information about the source of trades and the intention of the trader is likely to be communicated among dealers.

Preferencing arrangements may take the form of a stand-alone market maker who has formal or informal arrangements with institutional or retail brokers to channel order flow in certain stocks in which the market maker makes a market (purchased order flow). Alternatively, preferencing may occur when a vertically integrated market maker channels order flow from its brokerage division (internalized order flow). These two types of preferencing have different implications for the

[1] As pointed out by Lin, Sanger, and Booth (1998), there is no Chinese wall between market making and research divisions of NASDAQ firms, unlike the NYSE.

information content of the trade. Purchased order flow primarily originates from retail order takers and is characterized as being largely uninformed. From the market maker's point of view, purchased order flow allows the uniformed trades to be separated from the potentially informed trades, but it does not enable the market makers to ascertain the reason for a trade. Overall, purchased order flow does not reduce marketwide adverse selection, as informed traders still exist; it merely allows market makers to divide order flow into uninformed and potentially informed. Conversely, for internalized order flow, the reason for the trade is often known because of the information transfer that occurs between the broker and dealer.

We therefore assume that there are two types of investors: institutional and retail. Retail investors trade through retail channels such as brokerage houses and online sites whereas institutional traders trade directly through market makers or electronic communication networks (ECNs). The probability that a trade is retail is given by λ or institutional by $1 - \lambda$, where $0 \leq \lambda \leq 1$. Retail order flow is assumed to be uniformed and can be distinguished from institutional order flow because of its origin. The adverse-selection cost faced by market makers for trading with retail investors is designated θ_R.

Unlike retail investors, we assume that institutional investors may be informed and, further, that there may be communication about their information when the trade is handled through an internalization arrangement. In such a case, the market maker faces a lower adverse-selection cost than when the trade is placed directly. The adverse-selection cost for an institutional trade is θ_I, and θ_I^* for an internalized institutional trade. Our key assumption is that the adverse-selection costs are lower for internalized trades than for noninternalized trades, that is, $\theta_I^* < \theta_I$. This assumption is supported by others, such as Huang and Stoll (1996). Battalio, Jennings, and Selway (1999) suggest payment for order flow schedules may be related to the informational content of order flow, and Hagerty and McDonald (1996) argue that payment for order flow provides brokers with the incentive to separate informed and uninformed orders before routing them to the market. Thus, we can group trades into three groups: preferenced retail, institutional, and institutional internalized. **Q1**

We assume there are two constraints on the proportion of institutional trades that are internalized: the proportion of institutional investors that wish to use an internalization arrangement, which we designate as ϕ, and the number of market makers, m, relative to the number of institutions, n. This second constraint is realistic for two reasons. First, where the market maker is related to the investor (i.e., is within the same company) the market maker is unlikely to be able to credibly create other internalization arrangements with unrelated firms without a conflict of interest to the detriment of the outside firm. Second, even if independent market makers are able to enter into multiple internalization agreements with unrelated investors, the number of internalization agreements that are available depends on the number of market makers. For example, if there is only one market maker and 10 institutions, it is unlikely that there will be 10 internalization agreements. However, it is more reasonable to assume 10 internalization agreements if there are 10 market makers (i.e., one per institution). For simplicity, we assume that the proportion of institutions that can internalize their trades is $m/\phi n$. The probability that any trade is internalized is therefore: $(1-\lambda)(m/\phi n)$. This assumption is supported by empirical evidence from Chung, Chuwonganant, and McCormick (2004), who find a positive relation between the proportion of internalized volume and the number of market makers.

We can now estimate the adverse-selection cost for any trade based on the probabilities that the trade is retail or institutional, and if it is institutional, the probability that it is internalized. For retail, the cost is $\lambda\theta_R$; for noninternalized institutional, the cost is $[1-(m/\phi n)](1-\lambda)\theta_I$; and for internalized institutional, the cost is $(m/\phi n)(1-\lambda)\theta_I^*$. Thus, the total adverse-selection cost, *ASC,* for any trade ex ante is:

$$ASC = \lambda\theta_R + \left[1 - \frac{m}{\phi n}\right](1-\lambda)\theta_1 + \phi\frac{m}{\varphi n}(1-\lambda)\theta_I^*. \tag{1}$$

Taking the first derivative with respect to *m,* the number of market makers yields:

$$\frac{\partial ASC}{\partial m} = \frac{1}{\phi n}(1-\lambda)\left(\theta_I^* - \theta_I\right), \tag{2}$$

which is negative if $\theta_I^* < \theta_I$, that is, that the adverse-selection cost of an internalized trade is less than that of a noninternalized trade. Thus far we model *ASC* to be declining in the number of market makers; we now turn to the adverse-selection cost as a percentage of the spread. The overall spread comprises adverse-selection cost plus inventory and order-handling costs. We assume that the latter two are also affected by market maker competition, and greater competition will result in lower levels of these costs. However, there is a nonzero lower bound on the effect of competition on these costs. This lower bound does not exist for adverse-selection costs, as we assume that these can approach zero. For modeling simplicity, we assume that order processing and inventory costs, *OPINV,* have the following function form: $OPINV = \alpha + \beta/m$, where α and β are nonzero positive constants. Total spread is therefore $ASC + \alpha + \beta/m$. The adverse-selection cost as a percentage of the spread is:

$$\frac{ASC}{SPREAD} = \frac{\lambda\theta_R + \left[1 - \frac{m}{\phi n}\right](1-\lambda)\theta_I + \frac{m}{\phi n}(1-\lambda)\theta_I^*}{\lambda\theta_R + \left[1 - \frac{m}{\phi n}\right](1-\lambda)\theta_I + \frac{m}{\phi n}(1-\lambda)\theta_I^* + \alpha + \frac{\beta}{m}}. \tag{3}$$

Using the quotient rule to take first derivative with respect to *m,* the number of market makers, and rearranging yields:

$$\frac{\partial \dfrac{ASC}{SPREAD}}{\partial m} = \frac{\left[\alpha + 2\dfrac{\beta}{m}\right]\dfrac{1}{\phi n}(1-\lambda)\left(\theta_I^* - \theta_I\right) + [\lambda\theta_R + (1-\lambda)\theta_I]\dfrac{\beta}{m^2}}{\left[\lambda\theta_R + (1-\lambda)\theta_I + \dfrac{m}{\phi n}(1-\lambda)\left(\theta_I^* - \theta_I\right) + \alpha + \dfrac{\beta}{m}\right]^2}. \quad (4)$$

The sign of this derivative is unclear and reveals the potential nonlinearity between the percentage adverse-selection cost and number of market makers. The intuition behind this nonlinearity is straightforward. At low levels of market makers, the reduction in spreads due to increased market maker competition is greater than the reduction in adverse-selection costs due to internalization. As the number of market makers increases, the reduction in spreads due to competition declines and asymptotically converges to some constant, zero profit level we denote as α. Around this point, the decline in the adverse-selection costs outstrips the decline in overall spreads, and the adverse-selection cost as a percentage of the spread declines. The location of the inflexion point is an empirical question. However, in our sample, the minimum number of market makers, *m,* is 14, and at this level we believe it is safe to assume that an increase in the number of market makers is unlikely to result in significant further reductions in spread.[2]

We therefore hypothesize that an increase in market makers will result in lower adverse-selection costs through the internalization process. Greater numbers of market makers are also likely to produce a more competitive marketplace, and this alone may lead to narrower spreads and, consequently, lower adverse-selection costs. Indeed, several authors examine the relation between the number of market makers and the size of the spread, although we know of no prior study relating market makers to the adverse-selection component. The general consensus of these studies is that a greater number of market makers results in a more competitive market place and, hence, lower spreads. Our study controls for the effect of market maker competition on spreads by examining adverse-selection costs as a percentage of the spread. **Q2**

Direct studies of preferencing arrangements are hampered by the fact that details of these arrangements are not publicly available and are usually closely guarded. However, Chung, Chuwonganant, and McCormick (2004), using proprietary data from NASDAQ, find that the two types of order preferencing combined

[2]The assumption that for 13 market makers and above, *ASC/SPREAD* is declining in market makers is supported by a simple paramatization of the model. For example, assuming $\phi n = 250$, $(1-\lambda)\theta_I^* = 0$, $(1-\lambda)\theta_I = 5$, $\alpha = 20$, $\beta = 5$, yields a spread of about 20 to 30 cents, an adverse-selection cost of about 20% of the spread, and an inflexion point at eight market makers. These numbers are all broadly in line with those presented in the empirical section later in the article.

(purchased and internalized) make up more than 70% of all NASDAQ volume. They find that internalized order flow is the dominant order-flow type for large-volume stocks with high levels of institutional ownership. Such stocks tend to have more market makers than smaller, less frequently traded stocks. They find a positive relation between the proportion of internalized volume and the number of market makers and a negative relation between the proportion of purchased volume and the number of market makers.

Aside from market making relationships, NASDAQ stocks may have lower adverse-selection costs because they tend to have less institutional ownership and are generally smaller and therefore may have fewer informed traders. Schultz (2000) argues that dealers specialize in certain industries and that this specialization leads to an informational advantage. Dealers also tend to make a market in stocks in which their firms have investment banking relationships and, again, may have private information from the underwriting process. In a similar vein, Lin, Sanger, and Booth (1998) suggest that greater information gathering by NASDAQ market maker firms reduces adverse-selection costs but increases overall costs—an argument consistent with the lower adverse-selection costs and higher spreads on NASDAQ.

III. Data Sources, Sample Selection, and Variables

Data to decompose the spread into adverse-selection and other costs is obtained from the Trades and Quotes (TAQ) database. The TAQ data are also used to compute average price, trade size, and volume for the sample. All other data are obtained from the Center for Research in Security Prices (CRSP), Compustat, FirstCall, TFSD Shareworld, and NASDAQ. We examine April, May, and June 1999. We use this period to have the year-end financial statement data and recognizing the delay in the publication of annual reports. Because of the potential effect of tick size reduction on spreads, we examine the adverse-selection components during a period sufficiently removed from the first tick size change (May and June 1997) and before decimalization (early 2001). Our key objective is to ensure that our sample does not span a tick size reduction. Although our study is conducted pre-decimalization, we do not expect our results to be different in the post-decimalization era. Weston (2000) examines changes in adverse-selection components following the change in tick sizes from eighths to sixteenths and finds no significant change. Chung, Chuwonganant, and McCormick (2004) also find no significant change in preferencing follow decimalization.

We begin our sample by screening the data. We exclude financial service firms (Standard Industrial Classification (SIC) codes 6000–6999), regulated utilities (SIC codes 4800–4829 and 4910–4949), American Depositary Receipts (ADRS), real estate investment trusts (REITS), foreign firms, stocks with a price less than $3.00, and all stocks that undertake a stock split during the sample period.

After these screens, our initial NYSE sample comprises 856 companies and the NASDAQ sample comprises 866.

Creating a Matched Sample

To examine the impact of market makers on adverse selection, we first seek to replicate previous studies that find that adverse-selection costs are lower for NASDAQ stocks but spreads are higher compared with NYSE stocks. This comparison entails the creation of a matched sample. We match stocks on the basis of stock attributes that are strongly associated with spreads. However, several methods of matching stocks are by various researchers. The choice of matching criteria is discussed in Bessembinder (2003), who finds that matching on volume trading activity and price, as in Chung, Van Ness, and Van Ness (2001), produces results similar to those of Bessembinder (1999), who matches on firm size. We employ the method used by Chung, Van Ness, and Van Ness (2001) and match on price, number of trades, trade size, and return volatility. We also replicate our results using the matching procedure of Bessembinder (1999), who matches on market capitalization, and generate similar findings.

We measure share price by the mean value of the quote midpoint and return volatility by the standard deviation of returns calculated from the midpoint of the bid and ask prices. We measure trade size by the average transaction size during the study period.

To obtain a matching sample of NYSE and NASDAQ stocks, we first calculate the following score for each NYSE stock using our entire sample of NASDAQ stocks:

$$score = \sum_{i=1}^{4} \left[\frac{\left(Y_i^N - Y_i^T\right)}{\left(Y_i^N + Y_i^T\right)/2} \right]^2, \tag{5}$$

where Y_i represents one of the four stock attributes; T and N refer to NASDAQ and NYSE, respectively; and Σ denotes the summation over i = one to four attributes. For each NYSE stock, we then select the NASDAQ stock with the smallest score, not greater than one, with no replacement. This procedure results in 334 pairs of NYSE and NASDAQ stocks that are similar in price, number of trades, trade size, and return volatility. We are unable to obtain ownership data from TFSD Shareworld on 114 of the 334 matched NYSE and NASDAQ stocks. Our final sample has 220 matched NYSE and NASDAQ stocks.[3]

[3]We performed our tests that did not need ownership data again with the 334 stocks and find that our results are unchanged.

TABLE 1. Descriptive Statistics of Variable Used in the Matching Procedure.

Variable	Exchange	Mean	Std. Dev.	Min.	Max.
Panel A. Match Sample Variables					
Score	Both	0.3859	0.2751	0.0140	0.9971
Share price ($)	NYSE	31.87	19.54	4.79	107.57
	NASDAQ	24.38	13.58	4.27	95.62
Number of trades	NYSE	13250	16466	1229	166476
	NASDAQ	14398	21841	1335	228738
Trade size	NYSE	692.15	667.23	613	3916
	NASDAQ	527.38	337.44	461	2108
Volatility (as percentage)	NYSE	2.40	1.47	.21	9.21
	NASDAQ	2.57	1.68	.24	9.21
Panel B. Other Variables					
Number of analysts	NYSE	9.63	5.26	2	29
	NASDAQ	6.35	3.49	2	21
Number of institutional owners	NYSE	299	270.88	8	2030
	NASDAQ	71	45.41	6	242
Percentage of stock owned by institutions	NYSE	52.24%	22.51%	3.27%	97.86%
	NASDAQ	51.65%	21.34%	1.26%	100%
Average number of daily market makers	NASDAQ	41.20	19.95	13.94	142.88
Herfindahl index of market maker concentration	NASDAQ	1415.82	568.96	371.84	4387.32

Note: To obtain a matching sample of New York Stock Exchange (NYSE) and NASDAQ stocks, we first calculate the following score for each NYSE stock using our entire study sample of NASDAQ stocks:

$$\sum\left[\left(Y_i^N - Y_i^T\right)\Big/\left\{\left(Y_i^N + Y_i^T\right)\Big/2\right\}\right]^2,$$

where Y_i ($i = 1$ to 4) represents one of the four stock attributes (i.e., share price, number of trades, trade size, and return volatility); T and N refer to NASDAQ and NYSE, respectively; and Σ denotes the summation over $i = 1$ to 4. We then, for each NYSE stock, pick an NASDAQ stock with the smallest score. This procedure results in 220 pairs of NYSE and NASDAQ stocks that are similar in price, number of trades, trade size, and return volatility. We measure share price by the mean value of the midpoints of all quoted bid and ask prices, and trade size by the average size during the study period. The number of trades is the total number of transactions during the study period. We measure return volatility by the standard deviation of returns calculated from the midpoints of bid and ask prices. Number of analysts is the number of analysts issuing earnings forecasts as reported on First Call for the month before the component estimation period. We exclude firms with less than two analysts. Number of institutional owners and percentage of stock owned by institutions are collected from Security and Exchange (SEC) 13-F filings as reported on TFSD Shareworld. Average daily number of market makers is from NASDAQ and averaged over the component estimation period. The Herfindahl index of market maker concentration is the average over the component estimation period. These two variables are not applicable for NYSE stocks.

We report the summary statistics of our matched sample in Panel A of Table 1. The average price of our NYSE sample is $31.87 and the corresponding figure for our NASDAQ sample is $24.38. The average number of trades and trade size for the NYSE sample are 13,250 and 692.15, respectively, and the corresponding figures for the NASDAQ sample are 14,398 and 527.38. The mean values of the standard deviation of returns for our NYSE and NASDAQ stocks are 2.40%

and 2.57%, respectively. Overall, our matching sample of NASDAQ and NYSE is fairly close in terms of price, number of trades, trade size, and return volatility.

Variables That Influence Adverse Selection

Following Van Ness, Van Ness, and Warr (2001), we control for other factors that determine adverse-selection costs. We present these variables in summary form in Panel B of Table 1. For adverse-selection problems to exist there must be some volatility in the stock price. Random deviations from true value are needed to provide an informed trader the opportunity to capitalize on his or her private information. There must also be informed investors present who are able to trade on these deviations from true value and profit at the expense of the market maker. We compute and discuss variables that proxy for these factors. As our main hypothesis is that the number of market makers, or concentration of market maker activity, is an important determinant of information content for NASDAQ stocks, we also compute the following variables.

Volatility. To capture volatility in the true price of the stock, we include the standard deviation of the quote midpoint, *SDMID,* as a measure of intraday volatility. For return volatility we use the standard deviation of the daily stock return, *SIGR,* and the standard deviation of daily volume, *SIGVOL*.

Informed Trader Variables. Following Van Ness, Van Ness, and Warr (2001) and Brennan and Subrahmanyam (1995), we use three variables to measure the potential presence of informed traders. These variables are the number of analysts following the stock, the percentage of stock held by institutions, and the number of institutional owners.

Evidence exists for an inverse relation between analyst following and adverse selection. For example, Brennan and Subrahmanyam (1995) find greater analyst following reduces information asymmetries. However, there is also evidence of a positive relation as in Chung et al. (1995), who examine the effects of number of analysts following a firm on the overall size of the spread. Van Ness, Van Ness, and Warr (2001), using a simultaneous equation model, also find a positive relation between adverse selection and analyst following for most of the decomposition models they study. The existence of a positive relation implies that analysts follow stocks with the greatest benefits for their information.

In addition to analyst following, we use two institutional ownership variables, as in Brennan and Subrahmanyam (1995). *LINST* is the log of the number of institutional owners and *LPINST* is the log of the percentage of stock owned by institutions. A greater number of institutional owners is consistent with less private information, as many institutions compete with each other and the market maker to profit from their private information. Conversely, a higher proportion of institutional ownership (*LPINST*) is consistent with more private information, as *LPINST* may measure the presence of (a few) large block holders.

Other Control Variables. Market value of equity is an important determinant of the speed of adjustment of a stock price to new information, possibly because of greater awareness of investors of larger firms. Additionally, if investors face some fixed cost in information production, they will tend to follow larger stocks where they can take larger positions. To the extent that larger firms have more information surrounding them, we expect larger firms to have smaller adverse-selection components. Market value is an important control variable, given that we are including the number of market makers in our regression analyses.

Industry dummies are assigned for the following two digit SIC code categories: Mining 10–14; Construction and Manufacturing 15–39; Transportation and Public Utilities 40–49; Wholesale and Retail Trade 50–59; Finance, Insurance, and Real Estate 60–67; and Services 70–96.

Measures of Market Maker Competition. We seek to examine the impact of market maker competition on adverse-selection costs. We use the number of market makers and the Herfindahl index of market maker concentration by volume of trades as proxies for the level of market making competition. For NASDAQ stocks we compute the average daily number of market makers (*AVGMM*) reported from NASDAQ market maker volume data over the three-month component estimation period. Market maker concentration for stock i is computed using the Herfindahl index as follows:

$$HERF_i = \sum_{j=1}^{n} \left[\frac{100 V_{i,j}}{\sum_{j=1}^{n} V_{i,j}} \right]^2, \tag{6}$$

where $V_{i,j}$ is the volume of trades in stock i handled by market maker j.

Chung, Chuwonganant, and McCormick (2004) argue that the Herfindahl index may be a better measure of market maker competition, as the number of market makers may overstate the level of competition. For example, Shultz (2000) reports that for May 1997 through February 1998, the average number of market makers was greater than 10, but the average Herfindahl is greater than 2,500 (consistent with four market makers dividing volume equally).[4]

Summary statistics for these variables are presented in Panel B of Table 1. NASDAQ firms tend to have fewer analysts following (6.35 vs. 9.63 for the NYSE) and fewer institutional investors (71 vs. 299 for the NYSE). However, the percentage of institutional ownership is similar for both markets, at about 52%. NASDAQ

[4]A high Herfindahl is consistent with less competition. The Federal Trade Commission (FTC) considers a Herfindahl of 1,000 or less to be an unconcentrated industry. A Herfindahl of greater than 1,800 is considered concentrated. A Herfindahl of 10,000 indicates one market maker has 100% of the volume.

stocks have an average of 41 market makers, with a maximum of 143.[5] The mean Herfindahl index of market maker concentration for NASDAQ stocks is 1416, with a range from 372 to 4387, indicating substantial variation in market concentration across our sample of stocks.

Adverse-Selection Components

The adverse-selection component of the spread is calculated for each stock using the estimation procedures of Glosten and Harris (1988), George, Kaul, and Nimalendran (1991),[6] and Lin, Sanger, and Booth (1995). Our choice of models merits explanation. The models we select are not the most recent models presented in the literature. In particular, we choose not to focus on the models of Madhavan, Richardson, and Roomans (1997) and Huang and Stoll (1997). Both of these more recent models provide erratic component estimates, which greatly reduces the sample size (see Van Ness, Van Ness, and Warr 2001). However, we replicate our tests using the plausible observations generated by these models and find that the results are not quantitatively different from those of the three other models. For all the models, adverse-selection components are initially computed as a percentage of the spread. We also compute the adverse-selection cost of transacting, defined as the adverse-selection cost as a percentage of the stock price (as in Brennan and Subrahmanyam 1995).

Glosten and Harris (1988). Glosten and Harris (1988) present one of the first trade indicator regression models for spread decomposition. A unique characteristic of their model is that the adverse-selection component, Z_0, and the combined order-processing and inventory-holding component, C_0, are expressed as linear functions of transaction volume. The basic model can be represented by:

$$\Delta P_t = c_0 \Delta Q_t + c_1 \Delta Q_t V_t + z_0 Q_t + z_1 Q_t V_t + \varepsilon_t, \tag{7}$$

where the adverse-selection component is $Z_0 = 2(z_0 + z_1 V_t)$ and the order-processing/inventory-holding component is $C_0 = 2(c_0 + c_1 V_t)$. P_t is the observed transaction price at time *t*, V_t is the number of shares traded in the transaction at time *t*, and ε_t captures public information arrival and rounding error. Q_t is a trade indicator that is $+1$ if the transaction is buyer initiated and -1 if the transaction is seller initiated. Glosten and Harris do not have quote data; hence, they are unable

[5] We obtained market maker data directly from NASDAQ; however, the number of market makers is also reported by CRSP. There is a substantial difference between these data sources. For example, for our sample, CRSP reports that the average number of market makers is 19 with a maximum of 53. Our market maker computation is based only on participants that had nonzero volume in the stock. We include volume by ECNs, but exclusion of the ECNs has no effect on the overall results.

[6] Neal and Wheatley (1998) modify Glosten and Harris (1988) and George, Kaul, and Nimalendran (1991) to accommodate transactions data. We use Neal and Wheatley's modified estimation procedures.

to observe Q_t. Having both trade and quote data, we use the Lee and Ready (1991) procedure for trade classification. We use ordinary least squares (OLS) to obtain estimates for c_0, c_1, z_0, and z_1 for each stock in our sample.

The bid-ask spread in the Glosten and Harris (1988) model is the sum of the adverse-selection and order-processing/inventory-holding components. We use the average transaction volume for stock i in the following to obtain an estimate of the percentage adverse-selection component, for each stock:

$$Z_i = \frac{2(z_{0,i} + z_{1,i}\bar{V}_i)}{2(c_{0,i} + c_{1,i}\bar{V}_i) + 2(z_{0,i} + z_{1,i}\bar{V}_i)}. \tag{8}$$

George, Kaul, and Nimalendran (1991). George, Kaul, and Nimalendran (1991) allow expected returns to be serially dependent. The serial dependence has the same effect on both transaction returns and quote midpoint returns. Hence, the difference between the two returns filters out the serial dependence. The transaction return is:

$$TR_t = E_t + \pi(s_q/2)(Q_t - Q_{t-1}) + (1-\pi)(s_q/2)Q_t + U_t, \tag{9}$$

where E_t is the expected return from time $t-1$ to t; π and $(1-\pi)$ are the fractions of the spread due to order-processing costs and adverse-selection costs, respectively; s_q is the percentage bid-ask spread (assumed to be constant through time); Q_t is a $+1/-1$ buy-sell indicator; and U_t represents public information innovations.

George, Kaul, and Nimalendran (1991) assume the quote midpoint is measured immediately following the transaction at time t. We use a T subscript to preserve the timing distinction for the quote midpoint. The midpoint return is:

$$MR_T = E_T + (1-\pi)(s_q/2)Q_T + U_T. \tag{10}$$

Subtracting the midpoint return from the transaction return and multiplying by 2 yields:

$$2RD_t = \pi s_q(Q_t - Q_{t-1}) + V_t, \tag{11}$$

where $V_t = 2(E_t - E_T) + 2(U_t - U_T)$. Relaxing the assumption that s_q is constant and including an intercept yields:

$$2RD_t = \pi_0 + \pi_1 s_q(Q_t - Q_{t-1}) + V_t. \tag{12}$$

We use the Lee and Ready (1991) procedure to determine trade classification. We use OLS to estimate the order-processing component, π_1, and the adverse-selection component, $(1-\pi_1)$, for each stock in our sample.

Lin, Sanger, and Booth (1995). Lin, Sanger, and Booth (1995) develop a method of estimating the empirical components using the effective spread. They define the signed effective half-spread, z_t, as the transaction price at time t, P_t, minus the spread midpoint, Q_t. The signed effective half spread is negative for sell orders and positive for buy orders. To reflect possible adverse information revealed by the trade at time t, quote revisions of λz_t are added to both the bid and ask quotes. The proportion of the spread due to adverse information, λ, is bounded by 0 and 1. The dealer's gross profit as a fraction of the effective spread is defined as $\gamma = 1 - \lambda - \theta$, where θ reflects the extent of order persistence.

Because λ reflects the quote revision (in response to a trade) as a fraction of the effective spread, z_t, and because θ measures the pattern of order arrival, Lin, Sanger, and Booth (1995) model the following:

$$Q_{t+1} - Q_t = \lambda z_t + \varepsilon_{t+1}, \tag{13}$$

$$Z_{t+1} = \theta Z_t + \eta_{t+1}, \tag{14}$$

where the disturbance terms ε_{t+1} and η_{t+1} are assumed to be uncorrelated.

Following Lin, Sanger, and Booth, we use OLS to estimate the following equation to obtain the adverse-information component, λ, for each stock in our sample:

$$\Delta Q_{t+1} = \lambda z_t + e_{t+1}. \tag{15}$$

We use the logarithms of the transaction price and the quote midpoint to yield a continuously compounded rate of return for the dependent variable and a relative effective spread for the independent variable.

IV. Results and Analysis

Comparison of the Spread Models

In this section we compare the estimates of the adverse-selection components for the NASDAQ and NYSE. In Table 2, Panel A we report the mean, standard deviation, and percentile value for each of the three adverse-selection models and for the spreads for the matched set of NYSE and NASDAQ stocks. The three adverse-selection models yield plausible component estimates and a consistent finding—NASDAQ stocks have smaller adverse-selection components than NYSE stocks. The mean adverse-selection component for NYSE ranges from 39.9% of the spread (Glosten and Harris 1988) to 48.2% of the spread (George, Kaul, and Nimalendran 1991). For NASDAQ stocks the range is 9.7% (Glosten and Harris

TABLE 2. Descriptive Statistics of Adverse-Selection Components and Spreads and Tests of Differences.

Panel A. Descriptive Statistics of Adverse-Selection Components and Spreads.

					Percentile			
Variable	Exchange	Mean	Std. Dev.	Min.	25	50	75	Max.
GH	NYSE	0.3993	0.1292	0.0219	0.3160	0.4000	0.4873	0.7639
	NASDAQ	0.0971	0.0291	0.0338	0.0769	0.0955	0.1132	0.1901
GKN	NYSE	0.4815	0.0760	0.0041	0.4588	0.4942	0.5246	0.6197
	NASDAQ	0.2507	0.0545	0.0453	0.2210	0.2522	0.2849	0.3848
LSB	NYSE	0.4601	0.1473	0.1158	0.3487	0.4429	0.5504	0.9130
	NASDAQ	0.1684	0.0368	0.0772	0.1446	0.1661	0.1907	0.2782
Spread	NYSE	0.1548	0.0376	0.0677	0.1322	0.1511	0.1735	0.3375
	NASDAQ	0.2377	0.0972	0.0920	0.1643	0.2162	0.2911	0.5708
Percentage spread	NYSE	0.0068	0.0042	0.0011	0.0038	.00571	0.0085	0.0252
	NASDAQ	0.0118	0.0055	0.0014	0.0081	.01042	0.0146	0.0320

Panel B. Adverse-Selection Component Differences: NYSE vs. NASDAQ

	GH	GKN	LSB
Adverse selection expressed as percentage of spread	0.3022	0.2308	0.2917
	(38.50)***	(44.45)***	(29.90)***
Adverse selection expressed as a percentage of price	0.0015	0.0001	0.0010
	(15.654)***	(1.425)	(11.811)***

Panel C. Bid-Ask Spread Differences: NYSE vs. NASDAQ

Raw bid-ask spread	−0.0828
	(−15.411)***
Percentage bid-ask spread	−0.0049
	(−22.929)***

Note: Panel A reports the adverse-selection component as a percentage of the spread for the models of Glosten and Harris (1988, GH), George, Kaul, and Nimalendran (1991 GKN), and Lin, Sanger, and Booth (1995, LSB), and raw and percentage bid-asked spreads. Panel B reports *t*-tests for differences in adverse-selection components scaled by spread and price. Panel C reports *t*-tests for differences in the spreads for raw spread and spread scaled by price (percentage spread). In all cases there are 220 pairs of NYSE and NASDAQ stocks.

***Significant at the 1% level.

1988) to 25.1% (George, Kaul, and Nimalendran 1991). Both raw and percentage bid-ask spreads are narrower for NYSE stocks than for NASDAQ stocks.

In Table 2, Panel B we test for differences in the magnitudes of the adverse-selection components between the matched samples. We find that adverse-selection estimates for NYSE stocks are significantly larger than for similar NASDAQ stocks when adverse selection is expressed as a percentage of the spread. When expressed as a percentage of the price, the Glosten and Harris (1988) and Lin, Sanger, and Booth (1995) models are significantly smaller for NASDAQ stocks. The George, Kaul, and Nimalendran (1991) model is also smaller for NASDAQ stocks, although not at the 10% level. These results confirm the main findings of Affleck-Graves,

TABLE 3. Relation Between Market Makers and Herfindahl and Information Models.

Panel A. Relation Between Models and Market Makers and Herfindahl Index of Market Maker Concentration (NASDAQ Only)

	GH	GKN	LSB
Market makers	−0.3372***	−0.5682***	−0.1840***
Herfindahl	0.3470***	0.4981***	0.1548**

Panel B. Quartile Differences Related to Information Models (NASDAQ Only)

Variable	Quartile	GH	GKN	LSB
No. of market makers	Q1	0.1052	0.2775	0.1718
	Q4	0.0814	0.2005	0.1577
Difference	Q1–Q4	0.0237***	0.0770***	0.0141**
Herfindahl	Q1	0.0839	0.2106	0.1592
	Q4	0.1050	0.2785	0.1731
Difference	Q1–Q4	−0.0212***	−0.0678***	−0.0140**

Note: Panel A shows the correlations between the models and the number of market makers and the Herfindahl index of market maker concentration for NASDAQ stocks. Panel B reports the average adverse-selection component for the NASDAQ sample for each quartile based on number of market makers and Herfindahl index of market maker concentration. The models are Glosten and Harris (1988, GH), George, Kaul, and Nimalendran (1991, GKN), and Lin, Sanger, and Booth (1995, LSB). There are 55 observations per quartile. Quartile 1 is the lowest and quartile 4 is the highest.

***Significant at the 1% level.
**Significant at the 5% level.

Hegde, and Miller (1994). In Panel C, we report the results of paired *t*-tests for the spread and percentage spread of the sample. Both spread and percentage spread are significantly smaller for NYSE stocks. Thus, our data are consistent with previous work. We find adverse-selection components to be smaller for NASDAQ stocks, though spreads tend to be larger.

In Table 3, Panel A we report Spearman correlations between the adverse-selection components and the number of market makers and Herfindahl index for NASDAQ stocks. These data are also presented graphically in Figure I, which shows scatter plots of ln(Herfindahl) against ln(adverse selection/price). The adverse-selection components are negatively correlated with the number of market makers and positively correlated with the Herfindahl index of market maker concentration. In Table 3, Panel B we report the adverse-selection estimates for the NASDAQ stocks for quartiles based on the number of market makers, and the Herfindahl of market maker concentration. For all three adverse-selection models, the highest quartile of market makers has significantly less adverse selection than the lowest quartile of market makers. The opposite result holds for the Herfindahl measure. These results are consistent with a greater number of market makers leading to lower adverse-selection costs. In the next section, we examine the impact of market maker concentration on adverse selection in a multivariate framework.

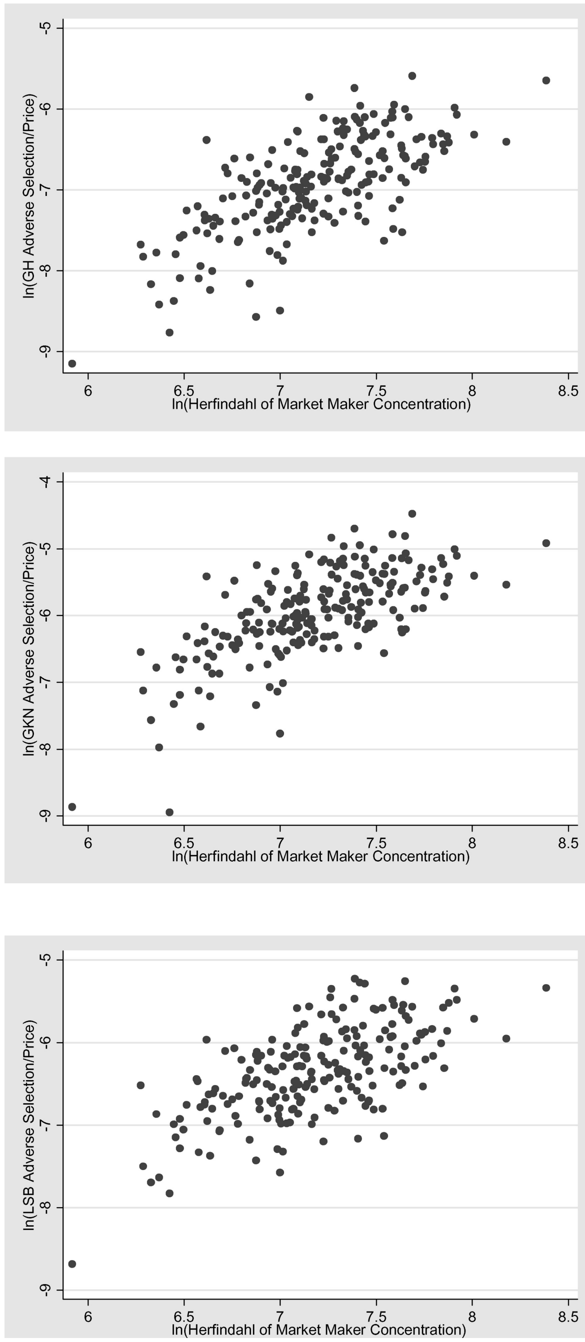

Figure I. Glosten and Harris (1988), George, Kaul, and Nimalendran (1991) and Lin, Sanger, and Booth (1995) Adverse Selection Component Estimates Plotted Against the Herfindahl Index of Market Maker Concentration.

Two-Stage Least Squares Regressions of Firm Characteristics on Adverse-Selection Components

We examine whether firm characteristics, other than the factors that determine spreads, explain the differences in adverse selection between our matched set of NYSE and NASDAQ stocks. We recognize the potential endogeneity problems that exist in our data and therefore use the method of Brennan and Subrahmanyam (1995). Specifically, greater analyst following may result in lower adverse selection (because of greater information production) or higher adverse selection (because analysts follow stocks where their potential gains from private information are greatest). Following Brennan and Subrahmanyam, we use two-stage least squares instrumental variable method (2SLS). The system is as follows and is modified from Brennan and Subrahmanyam's equations (18) and (19):

$$\begin{aligned} LAS = {} & \alpha_0 + \alpha_1 LANLYST + \alpha_2 LVOL + \alpha_3 LPRI + \alpha_4 LVAR + \alpha_5 LSIGR \\ & + \alpha_6 LSIGVOL + \alpha_7 LPINST + \alpha_8 LINST + \alpha_9 LNMVE \\ & + \alpha_{10} COMP + \varepsilon, \end{aligned} \tag{16}$$

Instruments for *LANLYST* are *LAS, LVAR, LPRI, LNMVE, LINST, LPINST,* IND_1, IND_2, IND_3, IND_4, where

$LAS =$ ln(adverse selection/price);
$LANLYST =$ ln(number of analysts following);
$LVOL =$ ln(volume);
$LPRI =$ ln(price);
$LVAR =$ ln(variance of spread midpoint);
$LSIGR =$ ln(standard deviation of returns);
$LSIGVOL =$ ln(standard deviation of daily volume);
$LNMVE =$ ln(market value of equity);
$LPINST =$ ln(percentage of institutional ownership);
$LINST =$ ln(number of institutional owners);
$IND_1, \ldots, IND_4 =$ industry dummies; and
$COMP =$ competition measure for NASDAQ stocks only—this is either ln(average number of market makers) or ln(Herfindahl of market maker concentration).

We modify the original Brennan and Subrahmanyam (1995) model to add the volatility measures (*LEVG, LSIGR,* and *LSIGVOL*) and informed trader proxies (*LINST* and *LPINST*). We also keep analysts in the model rather than transpose the analyst variable to the Admati and Pfleiderer (1988) informed trader model as done by Brennan and Subrahmanyam. Not transforming the analyst variable makes for a simplified interpretation, as our focus is not on the effect of analyst following.

The results for the adverse-selection models are presented in Table 4. For each spread decomposition model, we present the results of the 2SLS model for both the NYSE and NASDAQ firms. Hence, Table 4 should be read as three sets of regressions. Our null hypothesis is that there is no difference between the two exchanges; consequently, we expect all the coefficients from the NYSE and NASDAQ to be the same or similar for each adverse-selection model. However, we note that variations between the models for the same exchanges are possible. Of particular interest is the relation between market maker concentration and adverse selection.

Looking first at regression models (1) and (2) for all three adverse-selection models, we find that number of analysts (*LANLYST*) is positive and significant for both NASDAQ and NYSE stocks (although weakly for NASDAQ; George, Kaul, and Nimalendran 1991). This result is consistent with the hypothesis that analysts follow stocks where the benefit of their information production is the greatest. This result is the opposite of that found by Brennan and Subrahmanyam (1995), who find that analyst following reduces adverse-selection costs. We believe the differences in our results are due, in part, to the period studied. Brennan and Subrahmanyam examine 1988, and perhaps the role of analysts has changed since that time.

The other informed trader variables paint a similar picture. *LPINST* (the percentage of stock held by institutions) is positively related to adverse selection for NYSE stocks and for NASDAQ stocks (although weakly for the Lin, Sanger, and Booth 1995 model). Again, the magnitudes are similar for both the NASDAQ and NYSE stocks. *LINST* (the number of institutions owning the stock) is negatively related to adverse selection for all models for both NYSE and NASDAQ stocks. These variables suggest that large blockholders lead to higher adverse selection (as evidenced by the positive coefficient on *LPINST*). In addition, more institutional owners reducing adverse selection is consistent with numerous informed traders competing away their informational advantage.

The relation between the volatility measures and adverse selection are similar. *LVAR* (intraday quote volatility) is positively related to adverse selection for both markets although significant only for NASDAQ stocks. A similar pattern exists for *LSIGR* (interday return volatility) (although weakly for the Lin, Sanger, and Booth 1995 model for NASDAQ). As discussed earlier, higher volatility is necessary for informed traders to profit from their private information when the stock price deviates from true value.

We examine the impact of market maker concentration on the level of adverse selection of NASDAQ stocks in models (3) and (4) for each of the adverse-selection models. In model (3), we use the average daily number of market makers (*AVGMM*) as a proxy for market maker concentration. The results show that for all the models the number of market makers, *AVGMM,* is significantly and negatively

TABLE 4. Two-Stage Least-Squares Model of the Determinants of Dollar Adverse-Selection Component.

	GH				GKN				LSB			
	NYSE (1)	NASDAQ (2)	NASDAQ (3)	NASDAQ (4)	NYSE (1)	NASDAQ (2)	NASDAQ (3)	NASDAQ (4)	NYSE (1)	NASDAQ (2)	NASDAQ (3)	NASDAQ (4)
Intercept	6.684	5.207	2.831	−0.4214	4.837	5.535	3.759	0.6795	3.745	2.450	0.7674	−1.4808
	(8.77)***	(4.46)***	(1.94)	(−0.22)	(4.38)***	(6.67)***	(3.52)***	(0.47)	(4.54)***	(2.77)***	(0.68)	(−0.99)
LANLYST	0.280	0.972	1.244	1.2581	0.775	0.345	0.7184	0.7819	0.523	0.609	0.8832	0.8946
	(2.06)**	(3.44)***	(3.79)***	(3.91)***	(4.05)***	(1.68)	(2.82)***	(3.03)***	(3.59)***	(2.79)***	(3.31)***	(3.38)***
LVOL	−0.737	−0.633	−0.3679	−0.4803	−0.588	−0.529	−0.3206	−0.4183	−0.581	−0.502	−0.3090	−0.4044
	(−7.64)***	(−5.05)***	(−2.29)**	(−3.28)***	(−4.21)***	(−5.96)***	(−2.73)***	(−3.79)***	(−5.57)***	(−5.31)***	(−2.48)***	(−3.55)***
LPRI	−0.263	−0.596	−0.4670	−0.6315	−0.566	−0.675	−0.5898	−0.7411	−0.432	−0.633	−0.5479	−0.6738
	(−2.97)***	(−3.69)***	(−2.50)***	(−3.37)***	(−4.42)***	(−5.92)***	(−4.30)***	(−5.28)***	(−4.52)***	(−5.21)***	(−3.77)***	(−4.64)***
LVAR	0.045	0.148	0.1220	0.1508	0.062	0.109	0.0883	0.1146	0.029	0.074	0.0548	0.0772
	(1.43)	(3.08)***	(2.17)**	(2.70)***	(1.36)	(3.21)***	(2.15)**	(2.74)***	(0.84)	(2.04)**	(1.26)	(1.78)
LSIGR	0.610	0.174	0.1910	0.2142	0.383	0.163	0.1888	0.2145	0.363	0.082	0.1005	0.1179
	(5.60)***	(2.66)***	(2.53)***	(2.81)***	(2.42)**	(3.52)***	(3.39)***	(3.71)***	(3.07)***	(1.67)	(1.71)	(1.97)
LSIGVOL	0.158	−0.059	−0.0294	−0.1526	0.032	−0.067	−0.0405	−0.1539	0.157	0.081	0.1037	0.0125
	(2.04)**	(−0.55)	(−0.23)	(−1.19)	(0.28)	(−0.88)	(−0.44)	(−1.60)	(1.87)	(0.99)	(1.07)	(0.13)
LPINST	0.161	0.198	0.1834	0.2041	0.142	0.122	0.1206	0.1424	0.097	0.100	0.0939	0.1108
	(3.37)***	(2.45)**	(1.97)	(2.19)**	(2.05)**	(2.13)**	(1.77)	(2.03)***	(1.88)	(1.64)	(1.30)	(1.52)
LINST	−0.221	−0.345	−0.246	−0.2996	−0.163	−0.167	−0.0996	−0.1502	−0.174	−0.207	−0.1400	−0.1854
	(−2.94)***	(−3.51)***	(−2.14)**	(−2.67)***	(−1.49)	(−2.39)**	(−1.18)	(−1.77)	(−2.14)**	(−2.79)***	(−1.57)	(−2.12)**
LNMVE	0.124	0.098	0.1155	0.0932	0.106	0.053	0.0745	0.0563	0.049	0.021	0.0374	0.0210
	(2.43)**	(1.15)	(1.17)	(0.95)	(1.44)	(0.89)	(1.03)	(0.76)	(0.89)	(0.33)	(0.49)	(0.27)
LNAVGMM	–	–	−0.9839	–	–	–	−0.8669	–	–	–	−0.7596	–
			(−3.18)***				(−3.80)***				(−3.15)***	
LNHERF	–	–	–	0.5334	–	–	–	0.4983	–	–	–	0.3897
				(3.68)***				(4.53)***				(3.43)***

Note:

$$LAS = \alpha_0 + \alpha_1 LANLYST + \alpha_2 LVOL + \alpha_3 LPRI + \alpha_4 LVAR + \alpha_5 LSIGR + \alpha_6 LSIGVOL + \alpha_7 LPINST + \alpha_8 LINST + \alpha_9 LNMVE + \alpha_{10} LNAVGMM + \varepsilon.$$

Instruments for *LANLYST* are *LAS, LVAR, LPRI, LNMVE, LINST, LPINST,* IND_1, IND_2, IND_3, and IND_4, where *LAS* = ln(adverse selection/price), *LANLYST* = ln(number of analysts following), *LVOL* = ln(volume), *LPRI* = ln(price), *LVAR* = ln(variance of spread midpoint), *LSIGR* = ln(standard deviation of returns), *LSIGVOL* = ln(standard deviation of daily volume), *LNMVE* = ln(market value of equity), *LPINST* = ln(percentage of institutional ownership), *LINST* = ln(number of institutional owners), $IND_1, \ldots, IND_4$ = industry dummies, *LNAVGMM* = ln(average number of market makers) (NASDAQ stocks), and *LNHERF* = ln(Herfindahl index of market maker concentration) (NASDAQ stocks). The models are Glosten and Harris (1988, GH), George, Kaul, and Nimalendran (1991, GKN), and Lin, Sanger, and Booth (1995, LSB). Note that as R^2 statistics have no statistical meaning in the context of two-stage/instrumental variable models, we do not report them. In all regressions there are 220 pairs of NYSE and NASDAQ stocks.

***Significant at the 1% level.
**Significant at the 5% level.

related to the adverse-selection component for NASDAQ stocks.[7] Therefore, even after controlling for firm size and trading volume, NASDAQ stocks with more market makers tend to have less adverse selection. This result is consistent with the adverse-selection cost of trading being lower for internalized order flow where market makers "know their order flow," and internalized order flow being positively correlated with the number of market makers. The inclusion of the market maker variable also has little effect on the overall significance of the other variables in the system, indicating that *AVGMM* is not proxying for some other factor such as size or volume. Indeed, the inclusion of *AVGMM* has the most effect on the intercept. Model (4) uses *HERF* as a proxy for competition. Given that *AVGMM* may overstate the degree of competition between market makers, *HERF* may be a better measure for an individual stock. For all models we find that *HERF* is significantly and positively related to the level of adverse selection.

Table 4 regressions indicate that factors that determine adverse selection on the NYSE and NASDAQ are similar. However, the additional market maker variable is a significant determinant of adverse selection for NASDAQ stocks. The effect of *AVGMM* is economically significant; for example, the *AVGMM* coefficient in the Glosten and Harris (1988) model of −0.98 implies a 2.4% decline in adverse selection for each additional market maker over the average of 40. A move of 2 standard deviations in the number of market makers would translate into a 67% change in adverse selection. For the Glosten and Harris model, adverse selection on the NYSE is 40% of the spread compared with 10% on the NASDAQ. A 67% reduction in adverse selection for the NYSE stocks would lead to a value of about 13% adverse selection as a percentage of spread. Similar results are obtained for the George, Kaul, and Nimalendran (1991) and Lin, Sanger, and Booth (1995) models. The impact of market makers on the level of adverse selection is, therefore, economically significant and appears to go a long way in explaining the difference in adverse selection between NASDAQ and the NYSE. A similar level of economic significance is evident from examining *HERF*. A move from the minimum *HERF* (the most concentrated market) to the mean results in a decline of 71% for adverse selection as measured by the Glosten and Harris metric. Again, similar results are obtained from the George, Kaul, and Nimalendran and Lin, Sanger, and Booth models.

Although we establish that the dollar adverse-selection cost is related to the number of market makers, this result may be driven by the inverse relation between the number of market makers and the percentage bid-ask spreads.[8] Because

[7] We find similar results for *AVGMM* when we measure adverse selection using the Madhavan, Richardson, and Roomans (1997) and Huang and Stoll (1997) models. But as reported earlier, we do not focus on these models because of the high number of implausible estimates they generate.

[8] In unreported regressions using the model in equation (16), we find that percentage spreads are significantly and negatively related to the number of market makers and positively correlated with the Herfindahl index of market maker concentration.

TABLE 5. Two-Stage Least-Squares Model of the Determinants of Percentage Adverse-Selection Component.

	GH – NASDAQ			GKN – NASDAQ			LSB – NASDAQ		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	2.0809	0.8950	−0.7179	2.3667	1.7242	0.3552	−0.7011	−1.2396	−1.7999
	(2.40)**	(0.87)	(−0.53)	(4.59)***	(2.95)***	(0.45)	(−1.08)	(−1.69)	(−1.88)
LANLYST	0.7096	0.8286	0.8444	0.0658	0.2206	0.2700	0.3368	0.4085	0.4014
	(3.38)***	(3.52)***	(3.61)***	(0.51)	(1.56)	(1.89)	(2.09)**	(2.31)**	(2.30)**
LVOL	−0.2650	−0.1337	−0.1882	−0.1577	−0.0811	−0.1144	−0.1317	−0.0710	−0.1028
	(−2.84)***	(−1.18)	(−1.81)	(−2.86)***	(−1.26)	(−1.90)	(−1.90)	(−0.88)	(−1.40)
LPRI	−0.0021	0.0622	−0.0180	−0.0770	−0.0474	−0.1091	−0.0366	−0.0084	−0.0450
	(−0.02)	(0.47)	(−0.14)	(−1.09)	(−0.63)	(−1.42)	(−0.41)	(−0.09)	(−0.48)
LVAR	0.0719	0.0588	0.0729	0.0318	0.0243	0.0347	−0.0030	−0.0090	−0.0023
	(2.00)**	(1.48)	(1.84)	(1.51)	(1.08)	(1.52)	(−0.11)	(−0.32)	(−0.08)
LSIGR	0.0838	0.0908	0.1029	0.0706	0.0817	0.0942	−0.0092	−0.0047	−0.0007
	(1.72)	(1.70)	(1.89)	(2.46)**	(2.68)***	(2.98)***	(−0.26)	(−0.12)	(−0.02)
LSIGVOL	−0.0806	−0.0663	−0.1269	−0.0888	−0.0786	−0.1255	0.0589	0.0659	0.0404
	(−1.00)	(−0.75)	(−1.40)	(−1.88)	(−1.57)	(−2.40)**	(0.99)	(1.05)	(0.63)
LPINST	0.1323	0.1238	0.1345	0.0539	0.0547	0.0645	0.0326	0.0297	0.0344
	(2.19)**	(1.89)	(2.03)**	(1.52)	(1.46)	(1.68)	(0.73)	(0.64)	(0.74)
LINST	−0.2388	−0.1883	−0.2151	−0.0576	−0.0341	−0.0535	−0.0989	−0.0767	−0.0910
	(−3.26)***	(−2.33)**	(−2.69)***	(−1.33)	(−0.74)	(−1.15)	(−1.81)	(−1.33)	(−1.61)
LNMVE	0.1070	0.1148	0.1043	0.0612	0.0699	0.0634	0.0295	0.0339	0.0289
	(1.68)	(1.65)	(1.49)	(1.64)	(1.77)	(1.57)	(0.63)	(0.68)	(0.59)
LNAVGMM	–	−0.4780	–	–	−0.3290	–	–	−0.2307	–
		(−2.19)**			(−2.63)***			(−1.47)	
LNHERF	–	–	0.2637	–	–	0.2112	–	–	0.1059
			(2.55)***			(3.51)***			(1.45)

Note:

$$LAS = \alpha_0 + \alpha_1 LANLYST + \alpha_2 LVOL + \alpha_3 LPRI + \alpha_4 LVAR + \alpha_5 LSIGR + \alpha_6 LSIGVOL + \alpha_7 LPINST + \alpha_8 LINST + \alpha_9 LNMVE + \alpha_{10} LNAVGMM + \varepsilon.$$

Instruments for *LANLYST* are *LAS, LVAR, LPRI, LNMVE, LINST, LPINST,* IND_1, IND_2, IND_3, and IND_4, where *LAS* = ln(adverse selection/price), *LANLYST* = ln(number of analysts following), *LVOL* = ln(volume), *LPRI* = ln(price), *LVAR* = ln(variance of spread midpoint), *LSIGR* = ln(standard deviation of returns), *LSIGVOL* = ln(standard deviation of daily volume), *LNMVE* = ln(market value of equity), *LPINST* = ln(percentage of institutional ownership), *LINST* = ln(number of institutional owners), $IND_1, \ldots, IND_4$ = industry dummies, *LNAVGMM* = ln(average number of market makers) (NASDAQ stocks), and *LNHERF* = ln(Herfindahl index of market maker concentration) (NASDAQ stocks). The models are Glosten and Harris (1988, GH), George, Kaul, and Nimalendran (1991, GKN), and Lin, Sanger, and Booth (1995, LSB). Note that as R^2 statistics have no statistical meaning in the context of two-stage/instrumental variable models, we do not report them. In all regressions there are 220 pairs of NYSE and NASDAQ stocks.

***Significant at the 1% level.
**Significant at the 5% level.

we express adverse-selection cost as a percentage of the stock price, it should be expected to be lower for stocks with smaller spreads. Table 5 repeats the analysis (for only the NASDAQ stocks) using adverse selection as a percentage of the spread as the dependent variable. If the decline in dollar adverse selection is due to competition narrowing spreads, we expect the coefficients of the market maker variables to be insignificant. This is not the case. For the Glosten and Harris (1988) and George, Kaul, and Nimalendran (1991) adverse-selection components (expressed as a percentage of the spread), we find a negative relation with the number of market makers and a positive relation with the Herfindahl index of market maker concentration. For the Lin, Sanger, and Booth (1995) model, the signs on the market maker coefficients are similar but the significance is below normally acceptable levels. Although our main results are weakened by overall market maker competition, there remains a significant reduction in adverse selection for two of the three models.

V. Conclusion

We examine adverse-selection costs for a matched sample of NYSE and NASDAQ firms. Consistent with Affleck-Graves, Hegde, and Miller (1994), we find that adverse selection is higher on the NYSE than on NASDAQ. We hypothesize that the level of adverse selection on NASDAQ is due, in part, to the presence of multiple market makers and the ability of market makers to discern the reason for a trade that is routed through an internalized order flow agreement. Consistent with our hypothesis, we find that that adverse selection for NASDAQ stocks is negatively related to the number of NASDAQ market makers dealing in a stock (and positively related to the Herfindahl index of market maker concentration). Our results are economically significant and can explain a large portion of why NASDAQ stocks have less adverse selection than NYSE stocks.

Our results have implications for researchers using adverse-selection components as measures of the firm's asymmetric information. In particular, we find that adverse-selection models will likely understate the information problems for NASDAQ stocks versus NYSE stocks. This is not because NASDAQ stocks are more transparent, or because of a flaw in the adverse-selection models, but rather because adverse selection matters less to NASDAQ market makers than it does to NYSE specialists. Researchers using adverse selection as a measure of information opacity must be careful in direct comparisons of NASDAQ and NYSE stocks, as the institutional arrangements of the NASDAQ market reduce the adverse-selection risk that NASDAQ market makers face. Our findings also have policy implications, particularly in relation to the degree to which regulators should level the playing field with regard to the information advantages that may exist for a group of participants in a particular market. Preferencing and other arrangements may significantly protect NASDAQ market makers from the risk of trading with an informed trader but do little to help liquidity traders. Furthermore, it may be that the most informed traders on the NASDAQ are the market makers themselves.

References

Admati, A. and P. Pfleiderer, 1988, Selling and trading on information in financial markets, *American Economic Review* 78, 96–103.

Affleck-Graves, J., S. Hegde, and R. Miller, 1994, Trading mechanisms and the components of the bid-ask spread, *Journal of Finance* 49, 1471–88.

Barclay, M., 1997, Bid-ask spreads and the avoidance of odd-eighth quotes on NASDAQ: An examination of exchange listings, *Journal of Financial Economics* 45, 35–60.

Battalio, R., R. Jennings and J. Selway, 2001, The potential for clientele pricing when making markets in financial securities, *Journal of Financial Markets* 4, 85–112.

Bessembinder, H., 1999, Trade execution costs on Nasdaq and the NYSE: A post-reform comparison, *Journal of Financial and Quantitative Analysis* 34, 387–407.

Bessembinder, H., 2003, Issues in assessing trade execution costs, *Journal of Financial Markets* 6, 233–57.

Brennan, M. and A. Subrahmanyam, 1995, Investment analysis and price formation in securities markets, *Journal of Financial Economics* 38, 361–81.

Chung, K., C. Chuwonganant, and D. McCormick., 2004, Order preferencing, spreads, and dealer market share on NASDAQ: Evidence before and after decimalization, *Journal of Financial Economics* 71, 581–612.

Chung, K., T. McInish, R. Wood, and D. Wyhowski, 1995, Production of information, information asymmetry and the bid-ask spread: Empirical evidence from analysts' forecasts, *Journal of Banking and Finance* 19, 1025–46.

Chung, K., B. Van Ness, and R. Van Ness, 2001, Can the treatment of limit orders reconcile the differences in trading costs between NYSE and Nasdaq issues? *Journal of Financial and Quantitative Analysis* 36, 267–86.

Flannery, M., S. Kwan, and M. Nimalendran, 2004, Market evidence on the opaqueness of banking firm's assets, *Journal of Financial Economics*, 71, 419–60.

George, T. J., G. Kaul, and M. Nimalendran, 1991, Estimation of the bid-ask spread and its components: A new approach, *Review of Financial Studies* 4, 623–56.

Glosten, L. and L. Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123–42.

Hagerty, K. and R. McDonald, 1996, Brokerage, market fragmentation, and securities regulation in A. Lo, ed.: *The Industrial Organization and Regulation of the Securities Industry* (). **Q3**

Huang, R. and H. Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313–57.

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

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

Lin, J., G. Sanger, and G. Booth, 1995, Trade size and components of the bid-ask spread, *Review of Financial Studies* 8, 1153–83.

Lin, J., G. Sanger, and G. Booth, 1998, External information costs and the adverse selection problem: A comparison of NASDAQ and NYSE stocks, *International Review of Financial Analysis* 7, 113–36.

Madhavan, A., M. Richardson, and M. Roomans, 1997, Why do security prices change? A transaction-level analysis of NYSE stocks, *Review of Financial Studies* 10, 1035–64.

Neal, R. and S. Wheatley, 1998, Adverse selection and bid-ask spreads: Evidence from closed-end funds, *Journal of Financial Markets* 1, 121–49.

Petersen, M. and D. Fialkowski, 1994, Posted versus effective spreads: Good prices or bad quotes? *Journal of Financial Economics* 35, 269–92.

Schultz, P., 2000, The market for market making, Working paper, University of Notre Dame.

Van Ness, B., R. Van Ness, and R. Warr, 2001, How well do adverse selection components measure adverse selection? *Financial Management* 30, 77–98.

Weston, J., 2000, Competition on the Nasdaq and the impact of recent market reforms, *Journal of Finance* 55, 2565–98.

Queries

Q1 Author: Battalio, Jennings, and Selway (1999) is (2001) in the References. Please indicate which is correct.

Q2 Author: Please provide one or two citations as an example to support the following statement in the text: “Indeed, several authors examine the relation between the number of market makers and the size of the spread . . .”

Q3 Author: For Hagerty and McDonald (1996), please provide the name and location of the publisher and the page numbers of the article.