Chapter 7

The Impact of the Introduction of Index Securities on the Underlying Stocks: The Case of the Diamonds and the Dow 30

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

Richard S. Warr*
North Carolina State University, USA

We test the hypothesis that uniformed traders prefer to invest in a basket of stocks rather than a portfolio of individual stocks by examining the impact of the introduction of the Diamond Index securities on the underlying Dow 30 stocks. We find that following the introduction, the bid-ask spreads of the Dow 30 increase relative to spreads of matching stocks. However, we do not find a consistent change in the adverse selection components of the Dow stocks relative to the matching sample. Our overall results are consistent with either a movement of uninformed traders to the Diamonds or the increase of another component of the spread, such as inventory holding costs.

Keywords:

1. Introduction

The introduction of assets that trade baskets of securities has become increasingly common in recent years.[1] Subrahmanyam (1991) states that in a frictionless market these securities would be redundant, however the trading volume of index securities indicates that this is far from the case. In the presence of information asymmetries, these securities may provide liquidity traders with a low cost alternative to the direct investment in the underlying stocks. In this paper we examine the impact of the introduction of the Diamonds index securities on spreads and adverse selection.

In a market populated with both informed and uninformed traders, the uninformed typically bears the cost of trading against those more informed. A

*Corresponding author.

[1]For example, Diamonds (symbol — DIA) which track the Dow 30, SPDRs (symbol — SPY) which track the S&P 500 and Nasdaq 100 Trust (symbol — QQQ) which track the Nasdaq 100. DIA began trading on January 20, 1998 and QQQ began trading on March 10, 1999.

common characterization of this cost is the adverse selection component of the spread, but indirectly the cost is also reflected in the overall magnitude of the bid-ask spread. Kyle (1985) argues that the presence of traders who possess superior knowledge of the value of a stock can impose adverse selection costs on liquidity traders and market makers. Market makers are compensated for bearing this cost by widening the bid-ask spread, and ultimately recouping the cost from liquidity traders. For liquidity traders who wish to merely own a diversified portfolio there is no way to avoid these costs, as they must purchase each stock individually. Furthermore, these liquidity costs cannot be diversified away. However, the introduction of index securities provides liquidity traders with a vehicle for investing in a diversified portfolio without having to purchase individual securities. The adverse selection costs associated with index securities are likely to be significantly less than those for the underlying securities because the pooling of the stocks greatly reduces the ability of informed traders to profit from their stock-specific knowledge.

Subrahmanyam (1991) hypothesizes that upon the introduction of index securities, there should be an increase in the spread and the adverse selection component of the spread for the underlying stocks. These increases are caused by uninformed investors migrating to the new index securities, leaving a greater proportion of informed investors trading the underlying stocks. Because the market maker now faces a greater percentage of informed traders, he must increase the spread (or adverse selection component) to cover his cost of trading with more informed traders. Jegadeesh and Subrahmanyam (1993) examine the impact of the introduction of S&P 500 futures contracts on the spreads of the underlying stocks. They find that the spreads for a sample of S&P 500 stocks increase significantly following the introduction of the futures contract. They also find weak evidence that adverse selection components increased in the post futures period. Several drawbacks exist with the Jegadeesh and Subrahmanyam data and method. First, S&P 500 futures were originally issued in 1982, a time when only daily spread data is available; and second, their sample represents only a portion of the total 500 firms due to data constraints.

In this paper, we use broadly the same method of Jegadeesh and Subrahmanyam (1993) to examine the microstructure effects of the introduction of the Diamond Index securities which track the Dow Jones 30 Industrial Average. By computing spread and spread component data for all 30 firms and by creating a more representative control sample, we are better able to test the impact of the index stock on the underlying stocks. Additionally, we examine the overall trading costs of the Diamonds contract compared to the underlying basket of stocks. By doing so we hope to shed light on the relative costs of trading the basket versus trading the individual stocks.

Our results are mixed. The time period surrounding the introduction of the Diamonds is also one of market-wide decline in spreads. Such a decline makes it more difficult to discern the impact, if any, of the introduction of the Diamonds. However, after extensively controlling for factors that influence spreads, we find that, relative to matching stocks, the Dow 30 experiences a smaller decline in spreads around the introduction of the Diamonds. This result is consistent with uninformed traders moving from the Dow 30 to the Diamonds, and causing the market maker to increase spreads on the Dow 30 relative to other stocks.

A comparison of the adverse selection components of the Dow 30 stocks with the control sample reveals no significant impact of the Diamonds introduction. However, the power of such tests is weakened by the reliability of the adverse selection estimates, and our limited sample size.

The paper proceeds as follows. Section 2 discusses the introduction of the Diamonds, Section 3 discusses data issues, Section 4 presents our results and analyses and Section 5 concludes.

2. Diamond Index Securities

On January 20, 1998 the American Stock Exchange (AMEX) began trading Diamonds. Diamonds are a security that allows investors to buy or sell shares in an entire portfolio of the Dow Jones Industrial Average (DJIA) stocks, so investors can mimic the DJIA returns at a minimal cost (minimal when compared to purchasing each stock within the DJIA) by purchasing units (shares) in a trust consisting of DJIA stocks. Investors receive proportionate monthly cash distributions corresponding to the dividends that accrue to the DJIA stocks in the Diamonds portfolio, less trust expenses. The AMEX introduced this product to provide investors the advantages of indexing with the benefits of intraday trading, as unlike stock index mutual funds, Diamonds may be purchased throughout the trading day. The net asset value of Diamonds is computed each business day at the close of trading.

3. Data and Matching Portfolio

The data for this paper comes from the New York Stock Exchange TAQ (Trade and Quote) database and CRSP. To control for other factors that might be affecting spreads around the introduction of the Diamonds, we assemble a matching portfolio of stocks that represents our control group. To be eligible for matching, a stock must trade on the NYSE, not be in the Dow 30, and have data available on CRSP and TAQ for the study time period.

We match each stock in the Dow 30 with a NYSE counterpart on the basis of four stock attributes.[2] These attributes are share price, trade size, return volatility and market capitalization. Previous work has found the first three of these factors to be important determinants of the spread.[3] We also include market value as Dow stocks tend to be much larger than the average stock on the NYSE. The matching procedure uses data from the 30 trading days prior to the introduction of the Diamonds. We calculate the following composite match score (CMS) for each Dow stock in our sample with each of our selected match stocks:

$$\mathrm{CMS} = \sum_{k=1}^{4} \left[\frac{2\left(Y_k^{\mathrm{DOW}} - Y_k^{\mathrm{Match}}\right)}{\left(Y_k^{\mathrm{DOW}} + Y_k^{\mathrm{Match}}\right)} \right]^2,$$

where Y_k represents one of the four stock attributes, and the superscripts, Dow and Match, refer to Dow 30 stocks and potential match stocks, respectively. For each Dow stock, we pick the NYSE stock with the smallest score — as long as the score is less than 2. This matching procedure results in 30 pairs of NYSE stocks (the matching stocks are listed in the Appendix). Summary statistics of the Dow 30 and the matching portfolio are displayed in Table 1. Overall the quality of the match appears good. The notable outlier in the matching process was the market value of General Electric, however, this stock matched well on the other criteria.

4. Results and Analysis

4.1. *Spread, effective spread and price improvement*

We use three measures of trading costs in this study (percentage spread, traded spread and effective spread) and a measure of trading inside the spread (price improvement). Each of these measures is computed using transaction data and

[2] This procedure is similar to Huang and Stoll (1996) and Chung, Van Ness and Van Ness (2001).
[3] See Demsetz (1968), Benston and Hagerman (1974), Stoll (1978), McInish and Wood (1992), and Huang and Stoll (1996).

Table 1. Summary statistics for Dow 30 and matching stocks.

Time period is the 30 trading days before and the 30 trading days after the introduction of the Diamonds (January 20, 1998). All variables are measured daily. Score is the composite match score. A lower score indicates a better match. The score is computed using the following:

$$\mathrm{CMS} = \sum_{k=1}^{4} \left[\frac{2\left(Y_k^{\mathrm{DOW}} - Y_k^{\mathrm{Match}}\right)}{\left(Y_k^{\mathrm{DOW}} + Y_k^{\mathrm{Match}}\right)} \right]^2,$$

where Y_k represents one of the four stock attributes, and the superscripts, DOW and Match, refer to Dow 30 stocks and potential NYSE match stocks, respectively. Price is closing stock price. Volume is the average daily number of shares traded. MVE is the market value of equity, measured in thousands. Risk is standard deviation of daily returns. The full list of the Dow 30 stocks and their matches are presented in the Appendix.

		Mean	Median	Std Dev	Min	Max
Price	DOW 30	65.69	62.08	20.56	37.30	113.65
	Match	66.91	64.40	23.52	35.70	125.62
Volume	DOW 30	2,276,859	1,872,534	1,329,360	376,393	5,659,615
	Match	2,057,919	1,619,233	1,612,405	416,470	9,551,476
MVE	DOW 30	64,545,004	49,000,000	54,923,178	6,002,367	241,000,000
	Match	42,669,196	45,700,000	23,057,133	6,607,693	97,700,000
Risk	DOW 30	0.0185	0.0180	0.0030	0.0132	0.0251
	Match	0.0199	0.0187	0.0049	0.0104	0.0308
Score		0.3611	0.1406	0.4551	0.0169	1.6017

averaged for each security for each day of the study period. The percentage spread is calculated as:

$$\text{Percentage Spread}_i = \frac{(\text{Ask Price}_i - \text{Bid Price}_i)}{(\text{Ask Price}_i + \text{Bid Price}_i)/2)},$$

where Ask Price$_i$, is the posted ask price for stock i, and Bid Price$_i$, is the posted bid price for stock i, for each quote within the sample. As quotes occur both when trades occur and when they do not, we also calculate the spread that occurs when a trade occurs:

$$\text{Traded Spread}_i = (\text{Ask Price}_i - \text{Bid Price}_i).$$

To measure trading costs when trades occur at prices inside the posted bid and ask quotes, we calculate the effective spread using the following formula:

$$\text{Effective Spread}_i = 2|\text{Trade Price}_i - \text{Midpoint}_i|,$$

where Trade Price_i is the transaction price for security i and Midpoint_i is the midpoint of the most recently posted bid and ask quotes for security i. The effective spread measures the actual execution cost paid by the trader. Lastly, we measure the discount that is given during trading, namely, price improvement.

$$\text{Price Improvement} = (\text{Traded Spread} - \text{Effective Spread}).$$

Table 2 presents the means of the percentage spread and effective spread of the Dow 30 and the matching stocks for the 30 days before and 30 days after the introduction of the Diamonds. Surprisingly, for both the Dow and the matching stocks the effective spread and percentage spread declines in the

Table 2. Means tests of spread variables.

This table presents t-tests of the changes in the spread statistics for 30 trading days before and 30 trading days after the introduction of the Diamonds. The sample is the Dow 30 stocks and a sample of 30 matching stocks. The three spread measures are computed as:

$$\text{Percentage Spread}_i = \frac{(\text{Ask Price}_i - \text{Bid Price}_i)}{(\text{Ask Price}_i + \text{Bid Price}_i/2)},$$

$$\text{Effective Spread}_i = 2|\text{Trade Price}_i - \text{Midpoint}_i|,$$

$$\text{Traded Spread}_i = \text{Ask}_i - \text{Bid Price}_i.$$

	DOW 30	Matching Stocks	DOW-Match
Panel A: Effective Spread			
Before Introduction	0.0925	0.0985	−0.006
After Introduction	0.0903	0.0944	−0.004
Difference	0.0022	0.0041	−0.002
t-stat	2.787**	4.2379**	−1.836*
Panel B: Percentage Spread			
Before Introduction	0.0019	0.0020	−0.0001
After Introduction	0.0018	0.0018	−0.0000
Difference	0.0001	0.0002	−0.0000
t-stat	5.0838**	6.3238**	−1.4837
Panel C: Traded Spread			
Before Introduction	0.1175	0.1274	−0.0099
After Introduction	0.1138	0.1210	−0.0072
Difference	0.0037	0.0064	−0.0085
t-stat	2.9131**	4.2395**	−1.7016*

***Significant at the 1% level.
**Significant at the 5% level.
*Significant at the 10% level.

post-introduction period. The decline in spreads appears to be market wide and we consider it highly unlikely that it was caused by the introduction of the Diamonds, given the volume of the Diamonds relative to the market overall. This decline can be seen in Figure 1 which presents the average daily equally-weighted percentage spread for NYSE stocks for the time period under consideration. After the introduction on January 20, 1998, spreads are lower overall than in the prior period. The higher spreads around the end of December and early January may be due to the presence of several market holidays during this time. Chordia, Roll and Subrahmanyam (2001) find that market wide spreads tend to increase around holidays.

Referring to Table 2, the decline in spreads is smaller for the Dow 30 than the decline for the matching stocks. This result is consistent with spreads declining overall (due to some other unmeasured factor) but the decline in the spreads on the Dow 30 is lessened to some degree by the introduction of the Diamonds. An alternative explanation for this result is that the Dow stocks had spreads that were narrower than the matching firms prior to the Diamond's introduction and that some stocks were already trading close to their minimum tick size. During this time period the minimum tick size is 1/16th or 6.25 cents. The smallest average daily quoted spread was 7.42 cents (for General Electric), while the median quoted spread was 11.56 cents. Therefore, it is possible that for some of the most liquid stocks in the sample, the minimum tick provides a lower bound below which the quote cannot fall. However, most of the stocks in the sample have quotes that are substantially above the minimum tick both before and after the introduction of the Diamonds.

To control for other factors that might impact spreads, we employ the method of Chordia, Roll and Subrahmanyam (2001) who examine the impact of market wide and macroeconomic factors on market liquidity and trading activity. Chordia *et al.* find that the overall market return, the day of the week, holidays, and the change in the level of key interest rates significantly affect traded and effective spreads. We incorporate these variables into our regression analysis in Table 3 to control for other factors that might impact spreads around the introduction of the Diamonds. We combine the Dow 30 and matching firms into one sample and estimate the following regression model which allows us to observe the different slope coefficients for the Dow stocks compared to the matching firms.

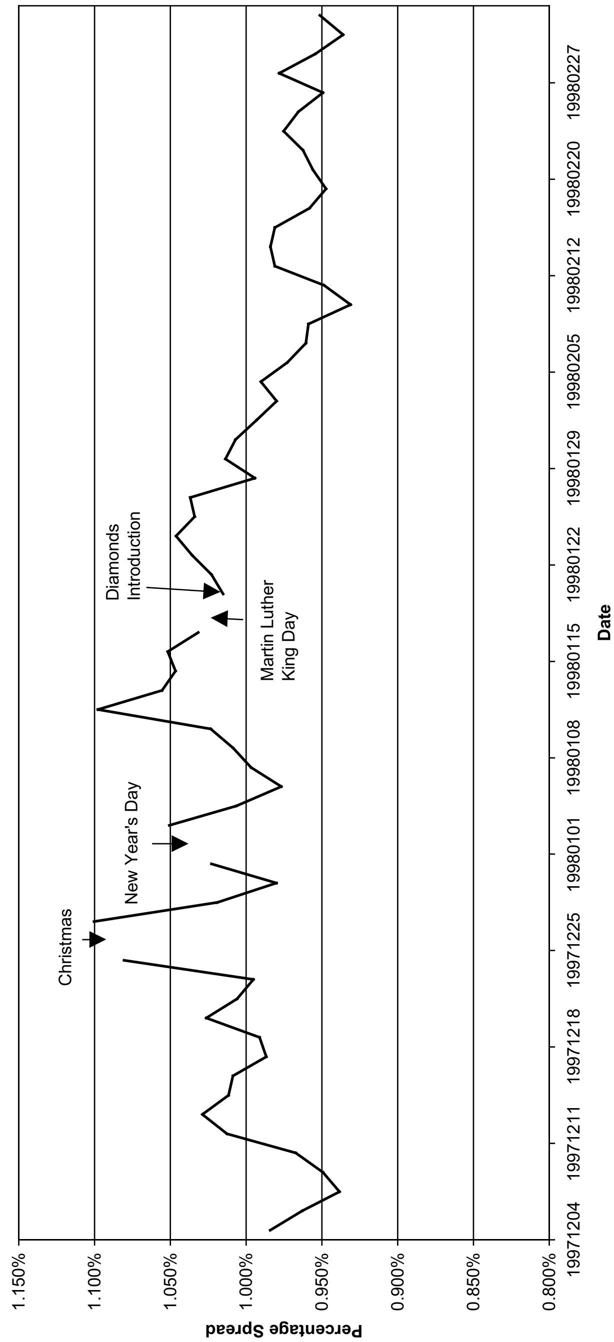

Figure 1. NYSE equally weighted daily quoted percentage spread.

Table 3. Feasible GLS estimates of the traded spread, effective spread and percentage spread.

FGLS is used to control for first order autocorrelation, heteroscedasticity, and cross sectional correlation. The data set is a panel of 30 Dow stocks and 30 match stocks. The time period is 30 days before and 30 days after the introduction of the Diamonds. The regression model is $[\text{Spread}_{it}] = a_0 + a_1\text{MD}_{it} + a_2\text{INTDUM}_{it} + b_2\text{MD}^*\text{INTDUM}_{it} + a_3\text{Price}_{it} + b_3\text{MD}^*\text{Price}_{it} + a_4\text{Trade Size}_{it} + b_4\text{MD}^*\text{Trade Size}_{it} + a_5\text{Trades}_{it} + b_5\text{MD}^*\text{Trades}_{it} + a_6\text{Sdmid}_{it} + b_6\text{MD}^*\text{Sdmid}_{it} + a_7\text{MKTUP}_t + b_7\text{MD}^*\text{MKTUP}_t + a_8\text{MKTDN}_t + b_8\text{MD}^*\text{MKTDN}_t + a_{9-12}\text{Day of the week Dummies}_t + b_{9-12}\text{MD}^*\text{DayoftheweekDummies}_t + a_{13}\text{Holiday}_t + b_{13}\text{MD}^*\text{Holiday}_t + a_{14}\text{Short Rate}_t + b_{14}\text{MD}^*\text{ShortRate}_t + a_{15}\text{Term Spread}_t + b_{15}\text{MD}^*\text{TermSpread}_t + a_{16}\text{Quality Spread}_t + b_{16}\text{MD}^*\text{QualitySpread}_t + \varepsilon_{it}$. In this regression, the matched sample is represented by the dummy variable MD which takes a value of 1 if the stock is a matching firm and zero otherwise. The interaction variables are presented in columns 1A, 2A, and 3A. The key variable of interest is INTDUM, a dummy variable, which takes the value of zero before the diamonds introduction and one after. The interaction term MD*INTDUM measures the differential change in spreads for the match firms relative to the Dow 30. Other control variables are: Price is the mean daily price for each stock, Trade Size is the mean daily trade size, Trades is the mean daily number of trades, SDMID is the standard deviation of the quote midpoint. MKTUP(DN) is the CRSP equally-weighted return if positive (negative) and zero otherwise. Monday, Tuesday, Wednesday and Friday are days of the week dummies. Holiday is a dummy if the trading day follows a holiday. Short rate is daily first difference in the Federal Funds rate, Term Spread is the daily change in the difference between the 10-year Treasury Bond and Short rate, Quality Spread is the daily change in the Moody's Baa or better corporate bond yield index and the yield on a ten year constant maturity Treasury Bond. Z statistics are in parenthesis. Each regression has 3,600 observations.

	Percentage Spread MD = 1		Effective Spread MD = 1		Traded Spread MD = 1	
	1	1A	2	2A	3	3A
Intercept	3.269 (422.66)**		76.602 (288.87)**		84.225 (192.55)**	
MD		−0.128 (−23.12)**		−9.232 (−21.05)**		−17.644 (−28.05)**

Table 3. (*Continued*).

	Percentage Spread MD = 1		Effective Spread MD = 1		Traded Spread MD = 1	
	1	1A	2	2A	3	3A
INTDUM	−0.043	−0.021	−1.922	−2.012	−2.661	−3.460
	(−36.10)**	(−13.94)**	(−19.08)**	(−10.78)**	(−9.72)**	(−11.60)**
Price	−0.018	0.001	0.281	0.120	0.418	0.233
	(−230.93)**	(21.90)**	(55.55)**	(19.06)**	(101.91)**	(33.15)**
Trade Size	−13.992	23.531	−962.042	1310.520	−1006.425	1798.619
	(−43.52)**	(56.96)**	(−56.71)**	(52.31)**	(−35.28)**	(47.78)**
Trades	−305.533	66.817	−9212.666	1942.798	−14783.888	4679.595
	(−206.05)**	(36.99)**	(−114.36)**	(16.36)**	(−123.74)**	(32.50)**
SDMID	0.769	−0.370	41.275	−19.390	62.658	−27.552
	(479.53)**	(−208.54)**	(240.32)**	(−108.90)**	(278.63)**	(−125.45)**
MKTUP	−6.695	−0.494	−279.064	5.083	−436.402	−49.022
	(−67.07)**	(−11.78)**	(−89.10)**	(0.60)	(−28.00)**	(−2.58)*
MKTDN	−1.263	−2.844	−25.714	−146.154	−41.538	−189.689
	(−11.05)**	(−63.02)**	(−7.92)**	(−15.78)**	(−2.41)*	(−9.07)**

Table 3. (*Continued*).

	Percentage Spread MD = 1		Effective Spread MD = 1		Traded Spread MD = 1	
	1	1A	2	2A	3	3A
Monday	0.011	0.026	0.595	0.411	1.210	1.001
	(5.04)**	(26.05)**	(10.95)**	(2.72)**	(4.60)**	(3.12)**
Tuesday	−0.009	0.026	−0.009	0.959	−0.036	1.943
	(−4.17)**	(28.13)**	(−0.17)	(6.70)**	(−0.15)	(6.56)**
Wednesday	−0.028	0.033	−0.508	0.809	−1.344	2.080
	(−13.81)**	(46.27)**	(−12.86)**	(6.53)**	(−6.47)**	(8.18)**
Friday	0.014	0.012	0.918	−0.441	1.815	−0.901
	(7.23)**	(16.19)**	(23.59)**	(−3.58)**	(8.73)**	(−3.56)**
Holiday	0.023	−0.005	0.749	0.198	1.309	0.682
	(20.02)**	(−7.15)**	(14.38)**	(1.62)	(5.46)**	(2.36)*
Short Rate	−0.488	−0.134	−17.478	−3.369	−32.919	−2.327
	(−37.77)**	(−24.80)**	(−46.13)**	(−3.04)**	(−16.01)**	(−0.93)
Term Spread	−0.426	−0.174	−15.446	−6.057	−29.112	−6.394
	(−32.69)**	(−31.96)**	(−40.12)**	(−5.38)**	(−13.90)**	(−2.50)*
Quality Spread	−0.761	−0.092	−17.735	−11.605	−43.146	−7.603
	(−22.87)**	(−7.22)**	(−20.60)**	(−4.27)**	(−8.88)**	(−1.28)
Wald Chi-Sqr	879,704		282,967		371,131	

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

$$
\begin{aligned}
[\text{Spread}_{it}] = {} & a_0 + a_1\text{MD}_{it} + a_2\text{INTDUM}_{it} + b_2\text{MD}^*\text{INTDUM}_{it} + a_3\text{Price}_{it} \\
& + b_3\text{MD}^*\text{Price}_{it} + a_4\text{Trade Size}_{it} + b_4\text{MD}^*\text{Trade Size}_{it} \\
& + a_5\text{Trades}_{it} + b_5\text{MD}^*\text{Trades}_{it} + a_6\text{Sdmid}_{it} + b_6\text{MD}^*\text{Sdmid}_{it} \\
& + a_7\text{MKTUP}_t + b_7\text{MD}^*\text{MKTUP}_t + a_8\text{MKTDN}_t \\
& + b_8\text{MD}^*\text{MKTDN}_t + a_{9-12}\text{Day of the week Dummies}_t \\
& + b_{9-12}\text{MD}^*\text{Day of the week Dummies}_t + a_{13}\text{Holiday}_t \\
& + b_{13}\text{MD}^*\text{Holiday}_t + a_{14}\text{Short Rate}_t + b_{14}\text{MD}^*\text{Short Rate}_t \\
& + a_{15}\text{Term Spread}_t + b_{15}\text{MD}^*\text{Term Spread}_t \\
& + a_{16}\text{Quality Spread}_t + b_{16}\text{MD}^*\text{Quality Spread}_t + \varepsilon_{it},
\end{aligned}
$$

where Spread_{it} is either the traded, effective or percentage spread. MD is an interaction dummy that takes the value of zero for the Dow stocks and one for the matching stocks. Price_{it} is the average trade price, Trade Size_{it} is the average trade size, Trades_{it} is the average number of trades, and Sdmid_{it} is the standard deviation of the quote midpoint. All of the variables are measured for $i = 1$ to 30 stocks on $t = 1$ to 60 trading days. INTDUM is an indicator variable that has the value of 0 on days before the Diamonds' introduction and 1 after the introduction. The remaining variables are those used by Chordia *et al.* (2001): $\text{MKTUP}_{t,}$, the CRSP equally-weighted daily return if positive and zero otherwise; $\text{MKTDN}_{t,}$ the CRSP equally-weighted return if negative and zero otherwise; days of the week dummies (Thursday is excluded); holiday dummies that take the value of 1 if the preceding day was a holiday; Short Rate_t, the change in the daily Federal Funds rate; Term Spread_t, the change in the difference between the 10-year Treasury rate and the Fed Funds rate; and Quality Spread_t, the change in the difference between the average yield on Moody's Baa rated corporate bonds and the 10-year Treasury rate.

Our data represents a balanced panel of 60 days with 60 observations per day. Such data will be subject to several econometric problems. Daily spreads are likely to be highly autocorrelated and heteroscedastic and there is the potential for cross-correlation in the panels. To control for these problems, we use Feasible Generalized Least Squares (FGLS) to estimate the regression models. By using FGLS, we control for autocorrelation, cross-correlation and heteroscedasticity.

The main results in Table 3 are contained in the coefficients of INTDUM (which measures the impact of the introduction on the Dow stocks) and the interaction between MD and INTDUM (which measures the marginal impact of the introduction on the matching stocks). In the table, we present each regression in pairs if columns, the first column of the pair (columns 1, 2, 3) being the Dow 30 stocks and the second column of the pair (columns 1A, 2A, 3A) being the matching firms' interactions, i.e., where MD = 1 in the main regression equation.

The dependent variable in the first regression (columns 1 and 1A) is the percentage spread. For both the Dow 30 and the matching sample, INTDUM is significantly negative, indicating that the introduction of the Diamonds reduces percentage spreads for both sets of stocks. In column 1A, the coefficient of INTDUM (the interaction of MD*INTDUM) is also negative and significant. This coefficient is important in our analysis as it presents the additional change in spreads for the matching firms. The total slope coefficient for the matching firms is the sum of the coefficients of INTDUM and MD*INTDUM. A negative MD*INTDUM indicates that while spreads decline for both the Dow 30 and the matching stocks, the decline is greater for the matching stocks. This result persists in columns 2A and 3A where the dependent variables are Effective Spread and Traded Spread, respectively.

Consistent with Chordia *et al.* (2001), we find that movements in the overall level of the market (captured by MKTUP and MKTDN) significantly impact the level of spreads. The change in the Fed Funds rate, the term spread and the quality spread are also significantly related to spreads for both the Dow 30 and matching stocks. Further, we find that holidays result in significantly higher spreads for both sets of stocks.

Overall, Table 3 shows that there is a significant reduction in spreads upon the introduction of the Diamonds and that this reduction is less for the Dow 30 than for the matching sample. This evidence is consistent with the hypothesis that uniformed traders in the Dow stocks migrate to the Diamonds, resulting in a relatively greater proportion of informed traders trading the Dow 30 stocks. The relative widening of spreads on the Dow 30 is also consistent with the market makers in those stocks anticipating an exodus of uninformed traders and widening spreads (relative to other stocks) to protect themselves accordingly. Our results could also be explained by an omitted variable problem, such as another unknown factor that could cause an impact on the Dow stocks around this time period. However, this factor would have to be correlated with Dow membership.

5. Adverse Selection Components

In this section we examine the impact of the Diamonds on the adverse selection components of the underlying Dow stocks and the matching sample. Following the introduction of the Diamonds, investors have the choice of two vehicles for investing directly in the Dow 30. If these investors are informed traders, they will trade the underlying stocks; however, if they are not informed, they should trade the Diamonds to avoid trading with the informed traders. The implication of this separation of traders is that the adverse selection costs for the underlying stocks should increase for the Dow 30 relative to the control group following the introduction of the Diamonds. We compute adverse selection components, using three different models,[4] for the 30 days before and for the 30 days after the introduction of the Diamonds. We use the models of Glosten and Harris (1988), George, Kaul and Nimalendran (1991) [both as modified by Neal and Wheatley (1998)], and Lin, Sanger and Booth (1995).

5.1. *Glosten and Harris (1988) (GH)*

Glosten and Harris 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:

$$\Delta P_t = c_0 \Delta Q_t + c_1 \Delta Q_t V_t + z_0 Q_t + z_1 Q_t V_t + \varepsilon_t,$$

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 did not have quote data, hence, they were unable to observe Q_t. Having both trade and quote data, we use the Lee and Ready (1991) procedure for trade classification. We use 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 model is the sum of the adverse selection and order processing/inventory holding components. We use

[4] See Clarke and Shastri (2000), Hegde and McDermott (2000), and Van Ness, Van Ness and Warr (2001) for a comparison of these and other adverse selection models.

the average transaction volume for stock i in the following to obtain an estimate of the percentage adverse selection component, for each stock:

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

5.2. *George, Kaul and Nimalendran (1991) (GKN)*

GKN allow expected returns to be serially dependent. The serial dependence has the same impact on both transaction returns and quote midpoint returns. Hence, the difference between the two returns filters out the serial dependence. The transaction return is:

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

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.

GKN assume the quote midpoint is measured immediately following the transaction at time t. As in Neal and Wheatley (1998), we will use an upper case T subscript to preserve the timing distinction for the quote midpoint. The midpoint return is:

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

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

$$2RD_t = \pi s_q(Q_t - Q_{t-1}) + V_t,$$
$$\text{where } V_t = 2(E_t - E_T) + 2(U_t - U_T).$$

Relaxing the assumption that s_q is constant and including an intercept yields:

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

As recommended by Neal and Wheatley, we use the Lee and Ready (1991) procedure to determine trade classification. We use OLS to estimate the adverse selection component, $(1-\pi_1)$, for each stock in our sample.

5.3. *Lin, Sanger and Booth (1995) (LSB)*

LSB develop a method of estimating empirical components of the effective spread following Huang and Stoll (1994), Lin (1993) and Stoll (1989). Huang and Stoll 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.

Since λ reflects the quote revision (in response to a trade) as a fraction of the effective spread z_t, and since θ measures the pattern of order arrival, Lin *et al.* model the following:

$$Q_{t+1} - Q_t = \lambda z_t + \varepsilon_{t+1},$$
$$Z_{t+1} = \theta Z_t + \eta_{t+1},$$

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

Following Lin *et al.*, we use OLS to estimate the following equation to obtain the adverse information component, λ, for each stock in our sample:

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

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 spread for the independent variable.

Table 4 shows the adverse selection measures for the 30 days before and after the initiation of the Diamonds on an equally-weighted basis. We measure adverse selection as a percentage of the spread and also, as a percentage of the price. The latter, "dollar" cost of adverse selection, is a better measure of the true cost of trading the stock as it controls for stock price and reflects the adverse selection cost based on the value of a trade rather than the number of shares traded.[5] All three models (both percentage and dollar) show a statistically significant decline in adverse selection for the Dow 30 (panel A) and for the matching stocks with the exception of dollar LSB (panel B) following the introduction of the Diamonds. We prefer to concentrate on the decline in dollar

[5] Dollar adverse selection is used in Brennan and Subrahmanyam (1995).

Table 4. Adverse selection component estimates for the Dow 30 and the matching stocks.

Adverse selection components are computed for 30 days before and 30 days after the introduction of the Diamonds. The component models used are Glosten and Harris — GH, George, Kaul and Nimalendran — GKN, and Lin, Sanger and Booth — LSB. Each panel presents adverse selection components computed as a percentage of the spread (%) and as a percentage of the stock price ($).

	Before	After	Difference	Two-Tailed T-Test	Sign Rank Test p-Value [Pos/Neg]
Panel A: Dow 30 Stocks					
GH %	0.2862	0.2729	−0.0133	−2.34**	0.035** [20/10]
GKN %	0.3982	0.3721	−0.0260	−3.57**	0.003** [21/9]
LSB %	0.4107	0.3872	−0.0236	−2.20**	0.079* [18/12]
GH $	0.000545	0.000487	−0.000058	−4.93**	<0.001** [27/3]
GKN $	0.0007739	0.0006761	−0.0000978	−5.98**	<0.000** [28/2]
LSB $	0.0007722	0.000684	−0.0000881	−4.39**	<0.001** [23/7]
Panel B: Matching Stocks					
GH %	0.3012	0.2728	−0.0283	−3.26**	<0.001** [26/4]
GKN %	0.4340	0.4082	−0.0256	−2.19**	0.013** [17/13]
LSB %	0.4043	0.3937	−0.0106	−0.59	0.766 [16/14]
GH $	0.0006025	0.0004992	−0.00010	−5.48**	<0.001** [26/4]
GKN $	0.0008804	0.0007465	−0.00013	−5.47**	<0.001** [27/3]
LSB $	0.0008105	0.0007353	−0.0000751	−2.21**	0.014** [21/9]
Panel C: Dow-Match					
GH %	−0.0150	−0.0001	0.0150	−1.48	0.131 [10/20]
GKN %	−0.0358	−0.0360	−0.0002	−0.02	0.586 [17/13]
LSB %	0.0064	−0.0065	−0.0130	−0.64	0.271 [21/9]
GH $	−0.0000575	−0.0000122	−0.0000453	−1.74*	0.050* [10/20]
GKN $	−0.0001065	−0.0000704	−0.0000362	−1.27	0.178 [12/18]
LSB $	−0.0000383	−0.0000513	0.000013	0.35	0.271 [21/9]

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

adverse selection rather than the decline in percentage adverse selection as the dollar measure captures both the decline in the component as a percentage of the spread and the decline in the overall spread. Panel C examines whether the change in adverse selection is different for the Dow 30 compared to the matching sample. All differences (except for the dollar LSB component) are negative, however, only one difference (GH dollar) is statistically significant. Therefore, we cannot conclude that the introduction of the Diamonds had an effect on the adverse selection costs of the underlying stocks. While this result does not support the overall spread effect, it is not surprising given the noisiness of the adverse selection models. An alternative explanation for our results is that some other factor, such as inventory costs, increases for Dow stocks relative to the matching stocks around the introduction of the Diamonds, thus increasing the spread relative to the matched stocks and offsetting the spread effect on adverse selection costs. A speculative explanation is that higher inventory costs could be the result of greater volatility in the Dow 30 stocks induced by increased index arbitrage following the introduction of the Diamonds.

6. Microstructure Characteristics of the Diamonds versus the Dow 30

In this section we examine the microstructure characteristics of the Diamonds compared to the Dow 30. Table 5 presents descriptive statistics of various quote and adverse selection measures. In panel A the Diamonds have lower adverse selection costs than the average of the Dow 30 stocks[6] for two out of the three models. Note that we cannot assign significance levels to these estimates as we only have one observation for the Diamonds and one average observation for the portfolio of the Dow 30 stocks. Panel A indicates that the various adverse selection models generate quite different estimates. Clarke and Shastri (2000) and Van Ness, Van Ness and Warr (2001) report that adverse selection models can produce widely different results for the same stocks.

Since the Diamonds represent a basket of stocks, we expect that its adverse selection would be small and close to zero since no informed trader would be able to profit on private information by trading the basket. A similar argument is made by Neal and Wheatley (1998), who find that adverse selection components for closed-end mutual funds are significantly greater than zero although they theorize that there should be little or no adverse selection for these securities. A possible explanation for the Diamonds having non-zero adverse selection is that informed traders can profit by trading with stale orders in markets where limit order traders do not update their orders continuously. Thus, even in a market where there should be no benefit to being informed about the underlying

[6]We use a price-weighted average, consistent with the construction of the Dow 30 index. Our results hold if we use an equally-weighted index.

Table 5. Microstructure statistics of the Dow 30 and the Diamonds.

The time period is 156 days after the introduction through August 1998. The composition of the Dow 30 changed in September 1998. The Dow variables are a price-weighted average of the component stocks of the Dow 30. Panel A presents the average adverse selection components as a percentage of the spread for the Dow stocks and the Diamonds. Note that there is only one estimate for group, therefore, statistical tests of differences cannot be undertaken. Panels B and C present quote and trade statistics for the Dow 30 portfolio and the Diamonds.

	DIA	Dow	Difference	Two-Tailed T-Test
Panel A: Adverse Selection				
GH	0.2025	0.2873	−0.0848	
GKN	0.5769	0.3822	0.1947	
LSB	0.1577	0.4023	−0.2446	
Panel B: Quote Statistics				
Spread	0.0959	0.1217	−0.0258	−12.93**
Traded Spread	0.1020	0.1136	−0.0116	−4.28**
Effective Spread	0.0691	0.0954	−0.0263	−17.33**
Price Improvement	0.0329	0.0181	0.0148	10.82**
Panel C: Trade Statistics				
Volume	585,148.7	1,923,138	−1,337,989	−41.14**
Trade Size	1,940.12	2,090.13	−150.01	−2.77**
Number of Trades	296.01	867.83	−571.83	−62.48**

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

security (such as the Diamonds), the adverse selection component of the spread may not be zero.[7]

The Dow 30 statistics shown in panels B and C are first calculated daily for each of the 30 stocks then averaged across the portfolio. Panel B shows the trading costs measures, and price improvement, for the Dow stocks and Diamonds. Diamonds have significantly lower spreads (0.0959) than the Dow 30 (0.1217). This implies that investors will have a cheaper round-trip transaction cost (approximately 2.5 cents) trading the Diamonds rather than the DJIA.[8] We find similar cost differences for traded spread (0.1020 for the Diamonds and 0.1136 for the Dow 30) and effective spreads. Additionally, we find that the amount of price improvement is larger for the Diamonds than for the Dow 30 (by approximately 1.5 cents). All of these findings are statistically significant indicating

[7]We would like to thank the referee for suggesting this explanation.
[8]These general results are robust when different trade sizes are examined.

that it is cheaper to trade the basket rather than the individual securities. Panel C presents the trade statistics for the sample, which show that the average volume of activity on a single stock in the Dow is greater than the total volume of the Diamonds. That the Diamonds are cheaper to trade yet have significantly lower volume than the Dow 30 stocks suggests that the order processing costs faced by the Diamond's market maker should not be lower than those faced by the individual stock market makers. Therefore the lower spreads of the Diamonds must be due to lower adverse selection or inventory costs. To the extent that making a market in the Diamonds exposes the market maker to less non-systematic risk than making a market in any single Dow stock, we would expect the Diamonds to have lower inventory costs as well.

In Table 6 we examine the factors that drive trading in the Diamonds securities. We proxy activity in two ways — trading volume (the number of shares

Table 6. Regression examining the causes of changes in the volume and number of trades of the Diamonds.

The time period is 156 days after the introduction of the Diamonds through August 1998. The dependent variables are the daily volume or the daily number of trades for the Diamond securities. DIA effective spread is the effective spread of the Diamonds. Dow 30 effective spread is a price-weighted average daily effective spread for the 30 Dow stocks. Volatility is the price-weighted average daily standard deviation of the quote midpoint return for the Dow 30. Volume is the price-weighted average daily volume of the Dow 30. Number of trades is the price-weighted average daily number of trades for the Dow 30. Newey West T-stats corrected for first order autocorrelation and heteroskedasticity are in parenthesis.

	DIA Volume	DIA Number of Trades
Intercept	−6.583	−2.609
	(−2.779)**	(−4.826)**
DIA Effective Spread	−6.298	−2.253
	(−3.280)**	(−4.827)**
Dow 30 Effective Spread	−30.112	−13.241
	(−3.062)**	(−4.640)**
Dow 30 Volatility	166.854	73.9222
	(4.466)**	(10.622)**
Dow 30 Volume	0.533	—
	(3.102)**	
Dow 30 Number of Trades	—	0.459
		(5.189)**
N	156	156
F(4, 151)	9.57	56.66

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

traded), and the number of trades. Our variables are computed for each day during our time period. We find that the trading volume of the Diamonds is positively related to the daily trading volume of the Dow 30 and also, the volatility of the Dow 30 (as measured as the standard deviation of the quote midpoint return). We also find that the number of trades per day of the Diamonds is positively related to the daily volatility of the Dow 30, and the number of trades in the Dow 30 stocks. These results indicate that the activity in the Diamonds moves in line with the overall activity in the underlying stocks.

7. Conclusion

We examine the impact of the introduction of the Diamonds stock index securities on the microstructure characteristics of the underlying Dow 30 stocks, and find that, when compared to a matched control group, the Dow 30 stocks exhibit a smaller decline in spreads. That spreads decline at all around the introduction of the Diamonds is puzzling; however, we attribute this decline to some other un-measured variable. However our tests prevent us from ruling out the explanation that as the market-wide liquidity improves, stocks with low liquidity improve more than those with high liquidity, and that the difference in liquidity improvement has nothing to do with the introduction of the Diamonds.

Adverse selection for both the Dow 30 and the control sample declines significantly upon the introduction of the Diamonds. However, the difference in the adverse selection components between the two groups is not statistically significant. While we believe that uniformed traders will migrate to the index security, and that this migration will result in higher trading costs on the underlying stocks [Subrahmanyam's (1991) hypothesis], we are not able to rule out the possibility that some other component, perhaps inventory costs, increases in relative terms for the Dow 30 upon the introduction of the Diamonds.

We find that while the Diamonds have, in general, lower adverse selection costs than the Dow 30, and that the adverse selection costs for the Diamonds are not trivial. This finding is surprising as we expect market makers in the Diamonds to face little risk from informed traders. A possible reason for the mixed adverse selection results is the poor empirical performance of adverse selection models in general.

We find that trading costs (spreads) are significantly lower for the Diamonds despite much lower volume. Additionally, Diamonds traders seem to get significantly more price improvement on their trades than do the traders in the Dow 30 stocks. Volume and trading activity in the Diamonds contracts is directly correlated with activity in the Dow 30 stocks as well as volatility of the Dow 30. Our results suggest that, for liquidity traders, the Diamonds contracts are a cheaper vehicle for achieving a diversified representation of the Dow 30 compared to buying the stocks directly.

8. Appendix: Dow Stocks and Matching Stocks

We match each stock in the Dow 30 with a NYSE counterpart on the basis of four stock attributes. These attributes are share price, trade size, return volatility, and market capitalization. Previous work has found the first three of these factors to be important determinants of the spread. We also include market value as Dow stocks tend to be much larger than the average stock on the NYSE. The data for matching comes from the 30 trading days prior to the introduction of the Diamonds (the matches are listed in the Appendix). We calculate the following composite match score (CMS) for each Dow stock in our sample with each of our selected match stocks:

$$\mathrm{CMS} = \sum_{k=1}^{4} \left[\frac{2\left(Y_k^{\mathrm{DOW}} - Y_k^{\mathrm{Match}}\right)}{\left(Y_k^{\mathrm{DOW}} + Y_k^{\mathrm{Match}}\right)} \right]^2,$$

where Y_k represents one of the four stock attributes, and the superscripts, Dow and Match, refer to Dow 30 stocks and potential match stocks, respectively. For each Dow stock, we pick the NYSE stock with the smallest score — as long as the score is less than 2. This matching procedure results in 30 pairs of NYSE stocks.

Dow Ticker	Matching Ticker	Composite Match Score
AA	BAX	0.1399
ALD	MTC	0.0446
AXP	ABT	0.1412
BA	PEP	0.1338
CAT	MDT	0.0819
CHV	MOB	0.0169
DD	LLY	0.0947
DIS	G	0.0469

Dow Ticker	Matching Ticker	Composite Match Score
EK	RD	0.0953
GE	PFE	0.9895
GM	SGP	0.1178
GT	HON	0.0395
HWP	F	0.1367
IBM	BMY	0.4423
IP	TMX	0.1013
JNJ	FNM	0.1801
JPM	FCN	0.2353
KO	BAC	1.3518
MCD	FTU	0.0898
MMM	WLA	0.1435
MO	MOT	1.4094
MRK	CCI	0.6670
PG	SBC	0.2654
S	PNU	0.0660
T	CPQ	0.9900
TRV	LU	0.3191
UK	TEN	0.0285
UTX	CL	0.0695
WMT	GTE	0.6936
XON	CMB	1.6017

References

Benston, G. and R. Hagerman, “Determinants of Bid-Asked Spreads in the Over-The-Counter Market.” *Journal of Financial Economics* 1, 353–364 (1974).

Brennan, M. and A. Subrahmanyam, “Investment Analysis and Price Formation in Securities Markets.” *Journal of Financial Economics* 38, 361–381 (1995).

Clarke, J. and K. Shastri, “On Information Asymmetry Metrics.” Working Paper, University of Pittsburgh (2000).

Chorida, T., R. Roll and A. Subrahmanyam, “Market Liquidity and Trading Activity.” *Journal of Finance* 56, 501–530 (2001).

Chung, K., B. Van Ness and R. Van Ness, “Can the Treatment of Limit Orders Reconcile the Differences in Trading Costs Between NYSE and NASDAQ Issues?” *Journal of Financial and Quantitative Analysis* 36, 267–286 (2001).

Demsetz, H., "The Cost of Transacting." *Quarterly Journal of Economics* 82, 33–53 (1968).

George, T. J., G. Kaul and M. Nimalendran, "Estimation of the Bid-Ask Spread and its Components: A New Approach." *Review of Financial Studies* 4, 623–656 (1991).

Hegde, S. and J. McDermott, "Firm Characteristics as Cross-Sectional Determinants of Adverse Selection." Working Paper, University of Connecticut (2000).

Huang, R. and H. Stoll, "Market Microstructure and Stock Return Predictions." *Review of Financial Studies* 7, 179–213 (1994).

Huang, R. and H. Stoll, "Dealer Versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE." *Journal of Financial Economics* 41, 313–357 (1996).

Jegadeesh, N. and A. Subrahmanyam, "Liquidity Effects of the Introduction of the S&P 500 Index Futures Contract of the Underlying Stocks." *Journal of Business* 66, 171–187 (1993).

Kyle, A., "Continuous Auctions and Insider Trading." *Econometrica* 53, 1315–1336 (1985).

Lee, C. and M. Ready, "Inferring Trade Direction from Intraday Data." *Journal of Finance* 46, 733–746 (1991).

Lin, J. C., "Order Persistence, Adverse Selection, and Gross Profits Earned by NYSE specialists." *Journal of Finance* July, 1108–1109 (1993).

Lin, L., G. Sanger and G. Booth, "Trade Size and Components of the Bid-Ask Spread." *Review of Financial Studies* 8, 1153–1183 (1995).

McInish, T. and R. Wood, "An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks." *Journal of Finance* 47, 753–764 (1992).

Neal, R. and S. Wheatley, "Adverse Selection and Bid-Ask Spreads: Evidence from Closed-End Funds." *Journal of Financial Markets* 1, 121–149 (1998).

Stoll, H., "The Pricing of Security Dealer Services: An Empirical Study of NASDAQ Stocks." *Journal of Finance* 33, 1153–1172 (1978).

Stoll, H., "Inferring Components of the Bid-Ask Spread: Theory and Empirical Tests." *Journal of Finance* 44, 115–134 (1989).

Subrahmanyam, A., "A Theory of Trading in Stock Index Futures." *Review of Financial Studies* 4, 17–51 (1991).

Van Ness, B., R. Van Ness and R. Warr, "How Well Do Adverse Selection Components Measure Adverse Selection?" *Financial Management* 30, 77–98 (2001).