How Effective Are Effective Spreads?
An Evaluation of Trade Side Classification Algorithms

Ananth Madhavan

Kewei Ming

Vesna Straser

Yingchuan Wang*

Current Version: November 20, 2002

Abstract

The validity of the most commonly used measure of market quality using public data, the effective spread, depends on the accuracy of trade side classification and the benchmark quote. Using a combination of NYSE's Trade and Quote Data (TAQ) and proprietary order data, we compare the performance and the consequent effective spread biases of two widely accepted algorithms for samples of Nasdaq and NYSE stocks in the post-decimalization environment. Results reveal much higher sensitivity of the effective spread estimates to the choice of benchmark quote in the current trading environment than previously reported. This is due to high discrepancy between the quotes recorded at the order submission and the order execution time resulting from high frequency of trade and quote updates. The magnitude of the bias is similar on both markets and can be reduced by lagging benchmark quotes. Overall our results demonstrate significant difficulty in the quest for the actual transaction cost measure and at the same time call for more detailed trade reporting/collecting rules.

* All authors are employed by ITG Inc., 380 Madison Avenue, New York, NY 10017.
The information contained herein is for informational purposes only and has been compiled from sources, which we deem reliable. ITG Inc. does not guarantee the accuracy or completeness or make any warranties regarding the results from usage. All information, terms and pricing set forth herein is indicative based upon, among other things, market conditions at the time of this writing and is subject to change without notice. Additional and supporting information is available upon request. ITG Inc. is a member of NASD, SIPC.
© 2002 ITG Inc. All rights reserved.
110502-90312

1 Introduction

Transaction costs and execution quality are topics of great interest from both a practical and academic viewpoint. Numerous studies have attempted to quantify differences in market quality across various market structures (Madhavan (2000) presents a survey). One of the key metrics for market quality is the effective spread, defined as the signed difference between trade price and the bid-ask midpoint prevailing at the time of order submission. Indeed, in January 2001, the U.S. Securities and Exchange Commission adopted Rule 11Ac1-5, which requires all market centers to publish their trading quality on a monthly basis. Such required trading quality measures include effective spread, the rate of price improvement, execution speed, and others. Accurate computation of effective spread requires two crucial pieces of information, the trade side and the prevailing quote. Commonly available transaction data sets, such as the NYSE's Trade and Quote (TAQ) data or Nastraq, provide neither, raising the possibility of serious biases in intermarket comparisons. As such the calculation of effective spread relies heavily on the accuracy of the underlined trade direction algorithm and the choice of the benchmark quote. Errors in either can result in significant overstatement (or understatement) of the actual effective spread and in turn lead to erroneous conclusions regarding relative performance of alternative market centers.

The objective of this paper is twofold. First, we want to evaluate the accuracy of the effective spread estimates calculated based on publicly available data as compared to the true spreads obtained from data that includes actual trade side and order submission time. Second, we want to assess the relative magnitude of biases due to errors in trade classification and due to incorrect assignment of the benchmark quote.

Several algorithms have been proposed to identify whether a trade was initiated as a buy or sell order. These algorithms include the quote rule, the tick rule, the Lee and Ready (LR) (1991) rule, and the latest addition, the Ellis, Michaely and O'Hara (EMO) (2000) rule. We chose to compare performance and consequent effective spread biases of the most frequently used algorithm, the LR rule, and its latest improvement, the EMO classification algorithm. Our analysis is conducted on samples of Nasdaq and NYSE listed stocks, respectively, trading in a decimal tick-size environment. We use a combination of NYSE's Trade and Quote Data (TAQ) and actual trade direction data proprietary of ITG Inc. (ITG) from June 2002. Effective spread and relative effective spread are computed based on true side data, LR-based side, and EMO-based side. We evaluate different levels of quote lags in applying the LR and EMO algorithms.

Evaluation of trade classification and consequent biases of the competing algorithms has already received much attention in recent financial literature.[1] However, these analyses were performed using data prior to the reductions in tick size beginning in 1997.[2] Previous research has shown that biases tend to increase with higher frequency of quote updates and trading volume and a smaller tick size. We revisit the issue as market conditions have changed considerably since previous analyses. In comparison to 1997, the average trading volumes have more than doubled and tick size has decreased to decimal increments.[3]

Indeed, our results demonstrate that effective spread estimates are increasingly dependent on the specifics of the algorithm used for their calculation. Specifically, in the current trading environment, the accuracy of the effective spread estimates is far more sensitive to the choice of the benchmark quote than previously believed. Analysis shows that this finding is due to high discrepancy between the quotes recorded at the order submission and order execution time, a consequence of high frequency of quote updates in a smaller tick environment. We find that the EMO algorithm based effective spread provides smaller bias resulting from the incorrect trade side classification. However, due to large differences between benchmark quote recorded at order submission and trade execution time, it is the LR algorithm based effective spread that is significantly closer to the true spread. When correcting for this time bias, we find that for NYSE stocks the

[1] They include Lee and Radhakrishna (2000), Odders-While (2000), Peterson and Sirri (2002), Finucane (2000), Ellis, Michaely and O'Hara (2000), and Piwowar and Wei (2001).

[2] The most recent analysis by Piwowar and Wei (2001) uses July 1999 Nastraq data. Since their data does not include the information about the order submission, their study is not directly comparable to ours. Peterson and Sirri (2002) study, which is the closest to our study in terms of data availability and questions addressed, uses transaction data from June 1997.

[3] For example, in 1997, trading volumes on NYSE and Nasdaq markets were 132.7 and 149.9 billion shares respectively. In 2001, the volumes jump to 301.0 billion shares on NYSE and 456.3 billion shares on Nasdaq. As of the end of September 2002, we have observed 263.7 and 314.5 billion shares on NYSE and Nasdaq markets.

smallest bias is achieved using EMO algorithm while for Nasdaq stocks the smallest bias is still achieved by LR algorithm.

While quote lags can effectively reduce the time delay bias, the actual reporting lag information is not readily available in public data such as TAQ or Nastraq. In turn, a recommendation on what is the optimal algorithm for accurate estimation of the effective spread becomes very difficult. Based on our data, we conclude that the optimal quote lags for the determination of trade side is 2 seconds for both algorithms. The optimal time lag for determination of the benchmark quote, however, is much more ambiguous. If using the LR method, no time lag in assignment of the benchmark quote results in the most accurate estimate for NYSE market while for Nasdaq stocks the best lag is 2 seconds. If the EMO algorithm is used, we recommend a quote lag of 10 seconds for NYSE stocks and 2 seconds for Nasdaq stocks. While the latter procedure provides the least amount of bias, the optimal trade delay may vary depending on the liquidity of the sample used in the analysis and may also change with time. We conclude that considerable caution should be used in the interpretation of the estimated effective spreads depending on the algorithm and time period of data used in the analysis.

The study proceeds as follows. In Section 2, we provide a brief description of trade classification algorithms and discuss how our study relates to previous work on the topic. In Section 3, we present sample selection and data considerations. In Section 4, we present and discuss the results. We conclude in Section 5.

2 Background and Previous Research

Four major trade classification algorithms are widely used: the quote rule, the tick test rule, the LR rule, and the EMO rule. While the quote method infers trade direction by comparing trade price to the prevailing quote, the tick test infers it by comparing the current trade price to the previous trade prices. Using the quote (tick test) rule, a transaction is a buy if the associated trade price is greater than the mid-quote (preceding price) and is a sell if the transaction is smaller than the mid-quote (preceding price). For the quote rule, a transaction is not classified if it is executed at the midpoint of the prevailing quote. The tick test, on the other hand, looks at the previous price change to classify a trade as buy or sell.

The LR algorithm combines quote and tick rules into the following algorithm. A trade is marked as a buy (sell) if the trading price is above (below) the prevailing quote midpoint. For trades with prices equal to the quote midpoint, the LR algorithm uses a tick test to determine its side. Upon observing that quote updates resulting from a trade were frequently recorded ahead of the trade, Lee and Ready (1991) introduced another feature to their algorithm. Instead of using the current quote to which the trade price is compared to determine the side of the trade, they propose to use the 5-second-lagged quote. At the time of their publication, Lee and Ready (1991) reported that their approach was superior to the then existing alternatives, which in turn contributed to highest frequency of usage of their algorithm compared to the tick test and the quote rule.

Following the research into the performance of the above algorithms for Nasdaq stocks, Ellis, Michaely and O'Hara (2000) introduce a new classification algorithm. Their approach represents a modification of the LR algorithm based on the observed poor performance of LR rule on off-quote trades. Ellis, Michaely and O’Hara (2000) propose that trades executed at the prevailing bid and ask quotes be marked as sells and buys. For all other trades (i.e., all trades inside or outside the prevailing bid-ask spread) a tick test is used to find the side of the trade. Therefore, the only difference between EMO and LR algorithms is the classification of orders executed at prices not equal to bid, ask and midpoint prices. By experimenting with different trade lags, they show that delaying the trade does not significantly improve accuracy of trade classification, nor the bias of the effective spread estimate.

The number of methods and their increased usage in empirical studies has prompted extensive research into the accuracy of these algorithms. Many of the existing studies such as Finucane (2000), Lee and Radhakrishna (2000), and Odders-While (2000), evaluate the accuracy of alternative trade classification algorithms using the NYSE's widely disseminated Trade, Order, Report, and Quote (TORQ) data. Given the nature of TORQ[4] data, these studies are distinct in their focus on the same sample of 144

---

[4] See Hasbrouck (1992) for a detailed discussion of the TORQ data.

NYSE listed stocks for a three-month trading period starting in November 1990. By providing information about the actual trade side, TORQ data presents a natural choice for the analysis at hand. However, the data is obsolete and may not represent the current trading conditions. While Lee and Radhakrishna (2000) only consider accuracy of LR algorithm, Finucane (2000) compares performance of LR and tick test methods, and Odders-White (2000) tests accuracy of the LR algorithm, the tick test, and the quote rules. Finucane (2000) also provides detailed analysis of situations in which algorithms are likely to fail and which factors are likely to affect algorithms' performance.

Using more recent data, a year of trading activity between September 1996 and 1997, Ellis, Michaely, and O'Hara (2000) test the accuracy and effective spread biases of LR, tick and quote rules for a sample of Nasdaq stocks. As mentioned earlier, they propose their own trade classification algorithm with a goal to improve upon existing algorithms in terms of accuracy and bias in the estimate of the effective spread. Their evaluation criteria, however, was based on the difference between quotes recorded at trade execution and trade reporting time rather than order submission time. Therefore, their results do not provide insight into the bias of their algorithm compared to the true effective spread.

A study by Peterson and Sirri (2002) compares the accuracy and biases of LR and EMO algorithms to the true effective spread based on the order submission time quotes and the actual order side. This study analyzed a sample of NYSE stocks for two separate two-week periods in June 1997. During the two sample periods in the study the minimum tick size changed from 1/8 to 1/16, which adds a new dimension in the comparison of classification methods. Based on the most complete data, they show that EMO algorithm produces better estimates of the effective spread than LR for NYSE stocks. Also, the accuracy is greatest and spreads are the least biased if quotes are not lagged relative to trades neither for trade direction algorithm nor for the assignment of the benchmark quote.

With much less complete but more recent data (July 1999 TAQ data), Piwowar and Wei (2001) find that in applying trade direction algorithms, quotes should be lagged by 2 seconds for both NYSE and Nasdaq stocks. Due to the lack of true side and the order submission time data, their results are based on an optimality criterion, which minimizes the proportion of small-sized trades executed outside of the quoted spread.

In general, the studies are in agreement regarding the accuracy of alternative trade classification algorithms for NYSE and Nasdaq stocks. The consensus is that the highest classification accuracy is achieved by EMO algorithm, followed by the LR method, the tick test, and the quote rule. The results concerning the bias of the effective spread are also consistent across the existing research and show significant overstatement of effective spread estimates by all methods. Specifically, for Nasdaq stocks, Ellis, Michaely and O'Hara (2000) report LR algorithm overstates the effective spread by 1.96%, while their own (EMO) algorithm overstates it by 1.34%. For NYSE stocks, Peterson and Sirri (2002) report LR algorithm overstates the effective spread on average by 8.8% when the quote size was 1/8 and 16.5% when quote size changed to 1/16. On the other hand, for the EMO algorithm, they report biases averaging at 6.6% for the 1/8 tick size and 6.1% for the 1/16 tick size. While different in methodology due to data availability, the studies indicate that overstatement of actual effective spread is much more dramatic for NYSE stocks than for Nasdaq stocks regardless of the side inference algorithm applied. Finucane (2000) and Peterson and Sirri (2002) have documented that biases of effective spread estimates are systematic in that they increase with the higher trading volume and higher frequency of quoting activity. In addition, the results in Peterson and Sirri (2002) also show that biases can increase as a result of a smaller tick size. Finally, previous research also documents that the biases are smaller when prices are benchmarked against quotes that are very close to the trade time for both trade signing and benchmark quote choice purposes.

Since these results were reported much has changed in the prevailing market conditions. Average trading volumes have doubled (see footnote 4), markets have switched to a decimal tick size, and the trading process became increasingly automated and faster. All of these changes suggest that the biases calculated on the current market data could be very different from those previously reported if algorithms are not adjusted to a different (faster) trading environment. At the same time, the number of market centers has also increased, adding to the importance and need for the accurate measures of trading costs for valid comparisons and analysis. We believe these factors warrant another look at the issue. Indeed, our results show a very different magnitude of bias than previously reported and suggest more difficulty in finding the optimal trade siding algorithm for both markets. In what follows, we re-assess the impact of these issues on the accuracy and the consequent effective spread bias of the two algorithms in greater detail. We first turn our attention to data considerations, followed by a discussion of the results.

3 Sample Selection and Data

The analysis is based on two samples each containing 96 stocks traded on the NYSE and Nasdaq respectively between June 1 and June 30, 2002. In order to achieve a valid comparison between the two markets, the stocks in the two samples are matched according to market capitalization and total trading volume during the period of observation. To obtain a representative sample, we include stocks of different levels of market capitalization. With this in mind, we first pull a stratified random sample of 96 NYSE listed stocks from the S&P 1500 Index. In the next step, we search for the corresponding Nasdaq stocks using a greedy algorithm in order to minimize pairwise difference in market capitalization and monthly trading volume. We end up with two samples of 96 stocks each, where the difference in market capitalization and trading volume is statistically insignificant.

Table 1 shows descriptive statistics sorted by market capitalization for each sample. It shows that we achieve a close match in terms of market capitalization and average monthly trading volume between the two samples. Despite these similarities in trading volume, we observe considerable differences between NYSE and Nasdaq traded stocks concerning average trading size and speed data. Specifically, reported average trading size for Nasdaq stocks is around 450 shares per trade across all capitalization quintiles. The average trading size for the NYSE listed stocks on the other hand increases with the market capitalization and ranges from 490 shares per trade for the least liquid stocks to 1560 shares for the most liquid stocks. A dramatic difference between the two markets is also observed in trading speed measured by average daily number of trades, quote updates and quote durations. The differences are particularly strong for very liquid stocks (quintiles 4 and 5), where Nasdaq stocks record 3 times as many quote updates and 3.5 times as many trades as NYSE listed stocks. Average quote duration on the Nasdaq market ranges from 138 seconds for least liquid stocks and only 3.6 seconds for the most liquid stocks. On the NYSE, quote durations are more evenly distributed among different liquidity quintiles and range from 50 seconds for lowest liquidity stocks to 11 seconds for the most liquid stocks. These differences are a result of different market structures prevailing on the two markets. On Nasdaq, automatic electronic execution via Electronic Communications Networks (ECNs) and Small Order Execution System (SOES) comprise almost half of the order flow, while on the NYSE, only orders up to 1099 shares receive automatic electronic execution (via Direct +). In turn, stocks are traded faster and quotes are updated more frequently on the Nasdaq market than on the NYSE. While these differences are significant and may suggest an improper match of the two market samples, they can not be eliminated because they are due to structural differences of the markets and not stock specific.

Transaction data used in the analysis comes from two sources: the NYSE maintained TAQ data and actual ITG proprietary order data. While TAQ data provides time-stamped trades and quotes data for all NYSE and Nasdaq traded stocks, it does not provide the information on whether a trade was initiated as a buy or a sell order. This information is available in a proprietary ITG database referred to herein as GLUE. GLUE contains information about all orders processed by ITG on behalf of its clients. In general, GLUE data records orders and trade prices and volumes, buy/sell indicator, order type, order-tracking identifier, order submission time, order execution time, and other information concerning the life of an order. Since GLUE data represents only a subset of TAQ trades, we combine the two data sets into a single data set in order to compare GLUE trades with the surrounding trades in TAQ.

Given the nature of GLUE, our study is therefore limited to a set of orders/trades handled by ITG. Even though ITG order flow consists primarily of institutional trades, we believe that ITG orders present a fairly representative sample of general order flow. We assert this claim based on the descriptive statistics reported in Table 1, which do not depart in any way from characteristics generally observed for NYSE and Nasdaq markets as a whole.

In our analysis, we only include those GLUE orders for which we can find an exact price and size match in the TAQ data and which were executed during regular trading hours. These include only market orders due to the fact that limit orders may execute at varying prices and are not considered as trade initiators in the existing literature. We also exclude odd-lot orders because on the NYSE these are executed automatically against specialist inventory at the prevailing quote and are not disseminated to the tape.[5] To be consistent, we also excluded odd-lot orders sent to the Nasdaq market. It is important to note that ITG clients have the option to change the order type before the order executes. Therefore, for more precise tracking of GLUE orders in concert with the TAQ data, all orders with changed order type status are also excluded from the final sample. On the TAQ side, trades with non-standard settlement conditions and trades reported on the regional stock exchanges are also excluded from the data.

When merging GLUE and TAQ databases, we implement the following approach. To find the matching GLUE trade in the TAQ database, we searched an interval of [+20, –5] seconds around the reported trade time in GLUE. The search was conducted in two stages, looking for the exact price and trade size match for each order. In the initial stage, we search for the first matching trade in the time interval [T, T + 20 seconds]. Here, T denotes GLUE reported trade time. If a match was not found, we search for the closest trade to the GLUE reported time in the time interval [T - 5 seconds, T]. Consistent with the previously reported speed of trading result (see Table 1), we observe a smaller time gap between order submission time and trade execution time on Nasdaq markets than on NYSE. Specifically, the median time gap for NYSE stocks is 11 seconds and only 2 seconds for the Nasdaq stocks. Our final sample contains 3047 matched orders for NYSE stocks and 4390 for Nasdaq stocks. The distribution of these orders by market capitalization quintiles is provided in Table 2. As expected, the distributions skew towards larger companies for both markets.

[5] For detailed discussion on order reporting see Hasbrouck, Sofianos, and Sosebee (1993).

4 Results

4.1 Accuracy of EMO and LR Classification Algorithms

As mentioned, the accuracy of a trade classification algorithm depends heavily on finding the optimal quote to which the trade is compared. Extensive previous research investigated the topic and suggested different quote delays to correct for potential trade reporting delays. Lee and Ready (1991) proposed that quotes should be lagged by 5 seconds. Ellis, Michaely, and O'Hara (2000) and Peterson and Sirri (2002) suggest no time lag in applying both the LR and the EMO algorithms. Piwowar and Wei (2001) show that a 2-second time lag should be used to achieve the smallest proportion of trades outside of the quoted spread. To find the optimal quote delay for trade signing in 2002 data, we also investigate the accuracy of LR and EMO algorithms for different quote lags for each market. These results are reported in Table 3.

Table 3 lists rates of successful classifications for LR and EMO with 0, 2, 5, and 10-second quote lags. We find that for both algorithms and for both markets, the highest accuracy is achieved by lagging quotes by 2 second. Applying a 2-second quote lag, accuracy of LR (EMO) is 89.51% (91.04%) and 96.90% (95.81%) for NYSE and Nasdaq stocks respectively. It is not surprising that our results are in agreement with Piwowar and Wei (2001) as they used more recent data from 1999, while the other studies utilized data from prior to and including 1997.

An interesting observation from Table 3 is that, with the 2-second quote lag, EMO has lower accuracy than LR for Nasdaq stocks. EMO correctly determines side for 95.8% of the sample, where LR achieves 96.9%. Ellis, Michaely, and O'Hara (2000) report that, with no quote lags, EMO marginally outperforms LR. Our results for 0-second quote lag scenarios confirm their conclusion with 93.8% accuracy for EMO comparing with 93.2% for LR. The 2-second quote lag changes the relative ordering of accuracy between the two algorithms.

Also, Table 3 shows that the accuracy for side classification algorithms is much higher for Nasdaq stocks than for NYSE stocks. We attribute this observation to differences in the distribution of trades relative to quotes observed on the two markets.

Namely, previous research has noted that accuracy of classification algorithms is highly dependent on the location of a trading price relative to the prevailing quote. Ellis, Michaely and O'Hara (2000) and Peterson and Sirri (2000) show that the most accurately classified trades are those executed at the quoted price while the least accurately classified trades are those executed inside the prevailing quote. Therefore, we should expect higher classification accuracy on the market with a higher proportion of trades executed at the quoted price.

Table 4 illustrates our results in this context. The table shows that a much higher concentration of Nasdaq trades is executed at the quote compared to NYSE trades. Specifically, 86% of Nasdaq trades are executed at the quote and only 67% of NYSE trades are executed at the quote. Consistent with the previous research, classification accuracy is highest for the 'at quote' category for both markets. This in turn contributes to better classification performance of both algorithms for Nasdaq stocks.

Nasdaq trades are more frequently executed at the quotes compared to NYSE trades due to the structural differences between the two markets. Automated execution, such as that prevails on the Nasdaq market (via ECNs and SOES), makes price improvement much more difficult than on a continuous auction market like the NYSE. This is because the nature of the electronic trading is to execute fast at the prevailing quote, though at the cost of less price improvement.[6] In addition, electronic execution implies that trades are reported immediately following execution.

In addition, Table 4 reports classification accuracy of the two algorithms sorted by firm size quintiles. Consistent with what is reported in the literature, companies with higher market capitalization have lower accuracy. For example, using the LR rule, accuracy for trades in top quintile Nasdaq stocks is 95.7%, while the figure is 100% for the bottom quintile. For the NYSE names, these numbers are 87.7% and 89.5% respectively. The same trend is observed for the EMO algorithm.

[6] Extensive discussion on trading differences between NYSE and Nasdaq (e.g., trading speed, price improvement) can be found in Battalio, Hatch, and Jennings (2000).

4.2 Bias in Measuring Effective Spread

If TAQ data is the only source of information, the researcher has to infer trade side from trading patterns observed close to the trade time. As mentioned in the introduction, Peterson and Sirri (2002), Finucane (2000), and Ellis, Michaely and O'Hara (2000) already thoroughly discussed biases introduced by various classification algorithms in computation of the effective spread. In this section, we provide new evidences on the topic and compare the magnitude of bias in NYSE and Nasdaq markets in a decimal tick size environment.

Microstructure researchers often use effective spread in measuring trading costs and execution quality of difference exchanges. The effective spread is defined to be

$$\text{Effective Spread} = 2 \times \text{D} \times (\text{Price - Midpoint})\,.$$

The binary variable D equals 1 for buy orders and -1 for sell orders. For one order, the effective spread is twice the midpoint benchmarked transaction cost. Studying effective spread figures for a large number of orders helps a researcher to evaluate average trading costs as well as execution quality. Effective spread measures transaction cost in dollars. To gauge relative transaction costs, researchers also use relative effective spread, defined by

$$\text{Relative Effective Spread} = \frac{\text{Effective Spread}}{\text{Price}}\,.$$

It is clear from the definitions that bias of the calculated effective spread comes from two factors. First, bias comes from the wrong determination of trading side. If a researcher is to use TAQ data only, the side of the orders has to be inferred by examining trading patterns close to the target trade time. The second bias comes from the errors in the choice of the benchmark quote, which should represent quote conditions at order submission time. We find that by applying the quotes at trade execution time creates large negative bias in effective spread calculations. This finding is in contrast to Peterson and Sirri (2000) who show that using trade time benchmark quotes causes serious positive bias in computing average effective spread. They report up to 17% bias in effective spread estimation using LR algorithm and up to 6% if using EMO algorithm.

While the results from Peterson and Sirri (2000) are helpful in understanding the consequences of applying a side-classification algorithm, their results are no longer representative for the current market environment. Their study period covers a time range when the minimum tick on NYSE was 1/8 and 1/16. The NYSE and Nasdaq changed to a decimal tick size in January and April of 2001, respectively. Moreover, they only reported bias in NYSE stocks. In this paper, we study both NYSE and Nasdaq markets in June 2002, and we observe very different results of the bias estimates.

According to Bacidore, Ross and Sofianos (1999), execution quality is best measured by benchmarking trades against prevailing quotes at the order submission time. Our study uses proprietary ITG order side data to compute the true effective spread. We want to compare the effective spread biases using LR and EMO with the highest attainable accuracy. Hence, we choose to apply a 2-second quote delay in implementing the side classification algorithms. We refer to them as optimal-LR and optimal-EMO algorithms henceforth.

Our analysis points to high sensitivity of the estimated effective figures due to differences in the choice of quote benchmarks to which the trade price is compared. This prompts us to calculate effective spread bias for different levels of quote lag. We compute the following four types of effective spread estimates:

a) true effective spread using the prevailing quotes at the order submission time and the true order side from ITG data;
b) trade-time effective spread using the true side, but the benchmark quote is the 0, 2, 5, 10 seconds lagged quote from the trade time;
c) LR effective spread using optimal-LR-inferred side with benchmark quotes lagged by 0, 2, 5, and 10 second from the trade time; and
d) EMO effective spread using optimal-EMO-inferred side with benchmark quotes lagged by 0, 2, 5, and 10 second from the trade time.

We compute these four measures of effective spread in order to evaluate two aforementioned components of the bias caused by the effective spread calculation. To evaluate the bias caused by the difference in benchmark quotes from the order submission to the trade time, we compare (b) and (a). To evaluate a portion of the bias coming from misclassification of trade side, we compare (c) and (d) to (a) for LR and EMO respectively. The results of these comparisons are listed in Tables 5 through 8.

Table 5 reports effective spread estimates using the above four approaches. Initially, we observe that the reported true effective spreads for two markets are consistent with reports in the literature. The table shows that the trading cost on a price only basis is lower on the NYSE comparing with the Nasdaq market. For example, for the top quintile NYSE (Nasdaq) stocks, true effective spread is estimated to be 2.979 (3.338) cents. This figure is computed as described in (a). The trade-time effective spread numbers are 2.587 cents, 2.772 cents, 2.876 cents, and 2.968 cents for 0, 2, 5, and 10 seconds quote lag respectively. Apparently, benchmarking trades to the prevailing quotes at the trade time is highly risky in that the effective spread is understated by 13%. Using the optimal-LR algorithm, we report effective spread estimates of 3.105 cents, 3.38 cents, 3.405 cents and 3.443 cents for 0, 2, 5, and 10 seconds quote lag respectively. These numbers are 2.682, 2.901, 3.028 and 3.117 for the optimal-EMO algorithm.

Table 6 analyzes the biases of the effective spread estimates reported in Table 5. We conduct three sets of comparisons. First, trade-time effective spread (b) is compared with the true effective spread (a) to evaluate the bias attributable to the change in midpoint from order submission time to the quote time. Second, LR effective spread (c) and EMO effective spread (d) are compared to the true effective spread (a) to evaluate the overall biases observed due to lack of actual side and order submission time data. Third, LR effective spread (c) and EMO effective spread (d) are compared to trade-time effective spread (b) to better gauge the bias causes solely by side misclassifications. The latter two comparisons are conducted in such a way that the trade time quote benchmark is lagged by the same number of seconds.

To illustrate, consider the top market capitalization quintile with 5-second quote lag as the benchmark quotes. We report -3.4% midpoint change bias for NYSE stocks and -1.7% for Nasdaq stocks. For NYSE stocks in this quintile, the optimal-LR algorithm, combined with 5-second quote delay when determining the benchmark quote, overstates the true effective spread by 14.3%, whereas the optimal-EMO based effective spread estimate is off by only 1.7%. The bias introduced by side misclassification is reported to be 18.4% and 5.3% for LR and EMO respectively. Clearly the overstatements are partially offset by the -3.4% bias caused by midpoint moves along the trade direction.

An immediate observation from Table 6 is that the midpoint change biases tend to diminish as longer quote lags were used. As the quote lag increases from 0 second to 10 seconds, the average midpoint change bias converges from -12% to -0.7% gradually for NYSE stocks. For Nasdaq stocks, the bias is at -2.1% for 2-second quote delay, and stayed roughly equal when the quote delay increases beyond 2-seconds.

Possible causes of the midpoint change biases come from fast current trading conditions. With the Nasdaq market increasingly automated, trade and quote updates are much more frequent than the previously covered study periods. In fact, median order duration (time from submission to execution) for Nasdaq stocks is only 2 seconds. In cases where there is a change of order in reporting, where a quote revision is reported earlier than the trade which causes it, the quote will more likely be closer to the trade price, which consequently reduces effective spread if used as the benchmark quote. For NYSE stocks, the median order duration is 11 second. This implies that a longer quote lag is needed to remove a similar portion of the midpoint change bias as for Nasdaq stocks. On the other hand, the NYSE is a continuous auction market. Specialists may disseminate the availability of the market order and adjust the quote before the order is recorded. Hence, the quote may move before the order is recorded as executed.

When comparing biases due to side misclassifications, EMO almost always dominates LR. For NYSE stocks, the bias from LR is always above 10%, where the maximum bias from EMO is only 1.4%. This is natural because, by definition, LR overstates the effective spread when the benchmark quote used in the side classification and the effective spread calculated are the same. This agrees with Peterson and Sirri (2002) assuming midpoint change bias is negligible during their study period.

The actual bias, which essentially is the sum of midpoint change bias and the side-misclassification bias, presents a much different picture. First, there is no universally optimal quote lag that should be applied across the board. For NYSE stocks, optimal-LR algorithm should be combined with no time lag in quote benchmark to achieve the closest match of the true effective spread. The overall bias using this setup is only -0.9%. However, for Nasdaq stocks, a researcher should apply 2, 5, or 10 seconds quote lags to achieve roughly 0.5% bias. For EMO it seems that the optimal time lag is 10 seconds for NYSE stocks and anything between 2 and 10 seconds for Nasdaq stocks. The fundamental reason for the mixed picture is that the actual bias is the result of the joint forces effect of the midpoint change bias and the side-misclassification bias. Although we have reasonable clarity in both the components of the overall bias, there is no clear optimal choice in the joint structure.

Tables 7 and 8 repeat the exercise for the relative effective spread. Based on Table 8, the LR algorithm should be combined with no quote lag for NYSE stocks, but with 5-second lag for Nasdaq stocks. For EMO, a researcher should use 5-second lag for NYSE stocks and anything between 2 to 10 seconds for Nasdaq stocks. This further strengthens the ambiguity of optimal quote lags. Comparing the magnitude of the effective and relative effective spread bias, we do not observe much difference, which is also different from Peterson and Sirri (2002).

Overall, our findings have two major implications. First, contrary to the existing literature, which always portraits EMO based effective spread as a better approach to measure trading quality, we have shown that there is another dimension in the biases of effective spread estimates, namely, the midpoint change bias. This bias is of large magnitude if no quote lag is used, which completely reverses the rank of performance of the two algorithms. For example, the actual bias for LR with no quote lag has an average bias of -0.9% comparing with the -11.9% reported for EMO. Second, it is rather hard to find the optimal algorithm in measuring execution cost by using only the pricing data. Although we have found the best combinations of side classification algorithms and quote lags, there is no guarantee that the structure will not change.

Consistent with the previous research, we also observe large differences between biases for the NYSE and Nasdaq markets even using the same algorithm and the same quote lag. These observed differences could have unwanted consequences when attempting to compare transaction cost on the two markets. Previously, we demonstrated that the accuracy of side classification rules for Nasdaq is noticeably higher than for NYSE. On the other hand, the magnitude of the midpoint change bias with no quote lags is roughly equal for both markets. Therefore, using a side classification algorithm and no benchmark quote lags in estimations could overstate the effective spread for one market while understate that of another. This in turn makes comparison between the two markets very difficult. For example, assume we want to use LR algorithm with no quote lags on stocks in the top market capitalization quintile. The actual effective spreads are 2.98 cents and 3.34 cents respectively for NYSE and Nasdaq stocks. However, the estimated effective spreads are 3.105 cents for NYSE stocks and 2.994 cents for Nasdaq stocks. The estimation procedure overstates the effective spread of NYSE stocks by 4.2% while understates that for Nasdaq stocks by 10.2%. Although the true effective spread for NYSE stocks is smaller than the Nasdaq stocks by around 10%, the estimated effective spread for the NYSE is actually higher. It clearly shows that considerable caution needs to be exercised when deciding on the type of estimation procedure and choice of benchmark quotes.

5 Summary and Conclusion

Using June 2002 transaction data, we analyzed the accuracy and subsequent effective spread biases of two trade classification algorithms for NYSE and Nasdaq markets. Effective spreads based on LR and EMO algorithms were compared to the actual effective spread calculated using actual trade side and quote prevailing at the time when the trade was submitted to the market. We decompose the bias into two parts; one due to incorrect trade classification and one due to incorrect benchmark quote assignment in the environment of high trading volume and decimal tick size.

We find several interesting results, all of which indicate one must use a considerable caution in choosing which algorithm to apply and what quote lag to use when attempting to measure the effective spread. We find that while EMO is usually more accurate than the LR algorithm in terms of trade classification, the overall bias of the two algorithms is dominated by the effect of the quote change between time of order submission and trade execution. Thus, while the ‘new’ optimal quote lag for trade side classification is 2 seconds for both markets and for both algorithms, the optimal lag for the assignment of the benchmark quote is much more ambiguous. It depends on the market under consideration and algorithm used to determine trading side. Analysis shows that this is due to the high discrepancy between the quotes recorded at the order submission and order execution time. The magnitude of this quote change bias can significantly influence which algorithm produces a better estimate of the actual effective spread. Based on our data, the best combination of the algorithm and the quote lag for the NYSE market is EMO algorithm with 10 seconds quote delay for benchmark quote determination. The best combination on the Nasdaq market is using LR algorithm with 2 seconds quote delay for the benchmark quote. Thus, while quotes no longer have to be lagged to correct for reporting delays, our results show that they should be delayed to correct for the quote change subsequent to the order submission.

In general, the optimal time lag for the benchmark quote assignment clearly corresponds to the average time needed to execute trades on respective markets. As this changes depending on the time of the data used in the analysis, market under consideration, and stocks' trading volume, our results should not be considered as an universal recommendation of which method to use to estimate the effective spread. Instead, our results should be understood as a testament to the difficulty of obtaining accurate calculations and a caution in interpretation of the effective spread estimates, particularly when applied to a comparison between different market centers.

6 References

Bacidore, Jeff, Ross, Katherine, and Sofianos, George (1999) “Quantifying Best Execution at the New York Stock Exchange: Market Orders” *NYSE Working Paper #99-05.*

Battalio, Robert H., Hatch, Brian, and Jennings, Robert (2000) “A Multidimensional Analysis of Retail Market Order Execution Quality” University of Notre Dame Working Paper.

Ellis, Katrina, Michaely, Roni, and O’Hara, Maureen (2000) “The Accuracy of Trade Classification Rules: Evidence from Nasdaq” *Journal of Financial and Quantitative Analysis* 35(4), 529-551.

Finucane, Thomas J. (2000) “A Direct Test of Methods for Inferring Trade Direction from Intra-Day Data” *Journal of Financial and Quantitative Analysis* 35(4), 553-576.

Hasbrouck, Joel (1992) "Using the TORQ Database," NYSE Working Paper.

Hasbrouck, Joel, Sofianos, Geroge, and Sosebee, Deborah (1993) “New York Stock Exchange Systems and Trading Procedures” NYSE Working Paper #93-01.

Lee, Charles M.C., and Ready, Mark J. “Inferring Trade Direction from Intraday Data” *The Journal of Finance* 46(2), 733-746.

Lee, Charles M.C., and Radhakrishna, Balkrishna (2000) “Inferring investor behavior: Evidence from TORQ data” *Journal of Financial Markets* 3, 88-111.

Madhavan, A. (2000), “Market Microstructure: A Survey” *Journal of Financial Markets* 3, 205-258.

Odders-White, Elizabeth R. (2000) “On the Occurrence and Consequences of Inaccurate Trade Classification” *Journal of Financial Markets* 3, 259-286.

Peterson, Mark, and Sirri, Erik (2002) “Evaluation of the biases in execution cost estimation using trade and quote data” mimeo.

Table 1
Descriptive Statistics

Market	Variable	Market Capitalization Quintile (5=top)				
		1	2	3	4	5
NYSE	Number of stocks	19	20	19	19	19
	Market capitalization (in million $)	359.44	1031.30	1912.01	3652.01	13736.39
	Trading Volume (in 1000 shares)	163.50	267.59	593.88	1031.95	3318.50
	Dollar trading volume (in 1000$)	2981.03	5982.24	17772.58	28394.77	106050.03
	Trade price	18.83	27.40	30.82	31.32	43.43
	Number of trades per day	215.74	374.46	683.02	814.51	1546.46
	Average trading size (in shares)	488.59	535.74	668.63	760.59	1561.43
	Number of quotes per day	635.37	969.11	1475.65	1619.14	2315.74
	Quote duration (in seconds)	50.42	26.99	17.71	15.95	11.19
NASDAQ	Number of stocks	19	21	16	23	17
	Market capitalization (in million $)	349.63	1035.50	1813.89	3529.52	13536.15
	Trading Volume (in 1000 shares)	161.47	284.97	600.38	1560.00	3527.04
	Dollar trading volume (in 1000$)	2194.45	8015.99	14159.16	40094.21	152122.99
	Trade price	16.94	28.63	24.91	32.90	42.60
	Number of trades per day	287.03	581.36	1087.49	2609.45	6427.87
	Average trading size (in shares)	419.20	486.45	471.97	456.43	414.27
	Number of quotes per day	572.21	1192.80	2016.82	4132.94	8337.66
	Quote duration (in seconds)	137.74	34.40	18.79	9.18	3.60

Table 2
Distribution of Market Orders by Market Capitalization Quintiles

Market	Market capitalization quintile				
	1 =Bottom	**2**	**3**	**4**	**5 =Top**
NYSE	6.50%	8.96%	16.44%	22.81%	45.29%
Nasdaq	4.85%	11.00%	9.00%	31.71%	43.44%

Table 3.
Accuracy of side inference algorithms at different quote lags

Market	Side Algorithm	Quote lag (in seconds)			
		0	**2**	**5**	**10**
NYSE	LR	87.7	**89.7**	89.2	88.0
	EMO	90.1	**91.1**	90.7	89.8
NASDAQ	LR	93.3	**97.1**	94.1	90.1
	EMO	93.9	**95.9**	92.9	89.9

Table 4.
Accuracy of side inference algorithms by company size and trade location

Market	Capitalization quintile	Trade Location	Percentage of trades by trade location	accuracy of LR	Accuracy of EMO
NYSE	1			89.90%	94.44%
	2			94.10%	93.36%
	3			89.22%	90.02%
	4			91.79%	92.94%
	5			87.97%	89.86%
		At mid-quote	3.64%	76.58%	76.58%
		At quote	67.57%	95.00%	95.00%
		Inside	25.40%	76.62%	82.95%
		Outside	3.38%	94.17%	90.29%
Nasdaq	1			100.00%	97.65%
	2			99.17%	98.55%
	3			97.96%	96.68%
	4			97.77%	96.25%
	5			95.86%	94.81%
		At mid-quote	0.68%	76.67%	76.67%
		At quote	86.26%	98.76%	98.76%
		Inside	5.40%	76.79%	58.65%
		Outside	7.65%	93.45%	90.18%

Table 5
Average Effective Spread by Market Capitalization Quintile

Bench-mark Quote Lag	Market Cap Quintile	True Side From Glue				LR Side		EMO Side	
		Order Submission Time Quote		Trade Time Quote					
		NYSE	Nasdaq	NYSE	NASDAQ	NYSE	NASDAQ	NYSE	NASDAQ
0	**1**	4.880	9.888	4.460	9.475	5.130	9.475	4.650	9.180
0	**2**	4.609	9.290	3.857	8.297	4.165	8.380	3.781	8.002
0	**3**	3.584	4.914	3.149	4.276	3.508	4.345	3.102	4.086
0	**4**	3.826	4.633	3.438	3.983	3.748	4.025	3.304	3.768
0	**5**	2.979	3.338	2.587	2.899	3.105	2.994	2.682	2.822
0	**Average**	**3.975**	**6.412**	**3.498**	**5.786**	**3.931**	**5.844**	**3.504**	**5.572**
2	**1**			4.510	9.982	5.300	9.982	4.840	9.659
2	**2**			4.158	9.150	4.511	9.250	4.112	8.846
2	**3**			3.365	4.892	3.743	5.038	3.310	4.802
2	**4**			3.616	4.478	3.986	4.598	3.516	4.303
2	**5**			2.772	3.134	3.380	3.382	2.901	3.148
2	**Average**			**3.684**	**6.327**	**4.184**	**6.450**	**3.736**	**6.152**
5	**1**			4.705	9.823	5.495	9.823	5.035	9.500
5	**2**			4.396	9.138	4.588	9.225	4.249	8.830
5	**3**			3.452	4.813	3.743	4.874	3.418	4.774
5	**4**			3.726	4.370	4.073	4.584	3.612	4.251
5	**5**			2.876	3.282	3.405	3.521	3.028	3.311
5	**Average**			**3.831**	**6.285**	**4.261**	**6.405**	**3.868**	**6.133**
10	**1**			4.810	9.837	5.460	9.837	5.140	9.514
10	**2**			4.505	9.099	4.794	9.144	4.413	8.782
10	**3**			3.612	4.669	3.890	4.720	3.578	4.675
10	**4**			3.812	4.384	4.096	4.690	3.732	4.349
10	**5**			2.968	3.322	3.443	3.563	3.117	3.356
10	**Average**			**3.942**	**6.262**	**4.337**	**6.391**	**3.996**	**6.135**

Table 6
Decomposition of Bias in Effective Spread Estimates

Quote Lag	Market Cap Quintile	Midpoint Change Bias		Actual Bias				Algorithm Bias			
				LR		EMO		LR		EMO	
		NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq
0	**1**	-8.6%	-4.2%	5.1%	-4.2%	-4.7%	-7.2%	15.0%	0.0%	4.3%	-3.1%
0	**2**	-16.3%	-10.7%	-9.6%	-9.8%	-18.0%	-13.9%	8.0%	1.0%	-2.0%	-3.6%
0	**3**	-12.1%	-13.0%	-2.1%	-11.6%	-13.4%	-16.8%	11.4%	1.6%	-1.5%	-4.4%
0	**4**	-10.1%	-14.0%	-2.1%	-13.1%	-13.7%	-18.7%	9.0%	1.0%	-3.9%	-5.4%
0	**5**	-13.1%	-13.1%	4.2%	-10.3%	-10.0%	-15.4%	20.0%	3.3%	3.7%	-2.6%
0	**Average**	**-12.1%**	**-11.0%**	**-0.9%**	**-9.8%**	**-11.9%**	**-14.4%**	**12.7%**	**1.4%**	**0.1%**	**-3.8%**
2	**1**	-7.6%	0.9%	8.6%	0.9%	-0.8%	-2.3%	17.5%	0.0%	7.3%	-3.2%
2	**2**	-9.8%	-1.5%	-2.1%	-0.4%	-10.8%	-4.8%	8.5%	1.1%	-1.1%	-3.3%
2	**3**	-6.1%	-0.4%	4.4%	2.5%	-7.6%	-2.3%	11.2%	3.0%	-1.6%	-1.8%
2	**4**	-5.5%	-3.3%	4.2%	-0.7%	-8.1%	-7.1%	10.2%	2.7%	-2.8%	-3.9%
2	**5**	-6.9%	-6.1%	13.5%	1.3%	-2.6%	-5.7%	21.9%	7.9%	4.7%	0.4%
2	**Average**	**-7.2%**	**-2.1%**	**5.7%**	**0.7%**	**-6.0%**	**-4.4%**	**13.9%**	**2.9%**	**1.3%**	**-2.4%**
5	**1**	-3.6%	-0.7%	12.6%	-0.7%	3.2%	-3.9%	16.8%	0.0%	7.0%	-3.3%
5	**2**	-4.6%	-1.6%	-0.4%	-0.7%	-7.8%	-4.9%	4.4%	1.0%	-3.3%	-3.4%
5	**3**	-3.7%	-2.0%	4.4%	-0.8%	-4.6%	-2.8%	8.4%	1.3%	-1.0%	-0.8%
5	**4**	-2.6%	-5.7%	6.4%	-1.0%	-5.6%	-8.2%	9.3%	4.9%	-3.1%	-2.7%
5	**5**	-3.4%	-1.7%	14.3%	5.5%	1.7%	-0.8%	18.4%	7.3%	5.3%	0.9%
5	**Average**	**-3.6%**	**-2.3%**	**7.5%**	**0.5%**	**-2.6%**	**-4.2%**	**11.5%**	**2.9%**	**1.0%**	**-1.9%**
10	**1**	-1.4%	-0.5%	11.9%	-0.5%	5.3%	-3.8%	13.5%	0.0%	6.9%	-3.3%
10	**2**	-2.2%	-2.1%	4.0%	-1.6%	-4.3%	-5.5%	6.4%	0.5%	-2.1%	-3.5%
10	**3**	0.8%	-5.0%	8.6%	-3.9%	-0.2%	-4.9%	7.7%	1.1%	-0.9%	0.1%
10	**4**	-0.4%	-5.4%	7.0%	1.2%	-2.5%	-6.1%	7.4%	7.0%	-2.1%	-0.8%
10	**5**	-0.3%	-0.5%	15.6%	6.8%	4.7%	0.5%	16.0%	7.3%	5.0%	1.0%
10	**Average**	**-0.7%**	**-2.7%**	**9.4%**	**0.4%**	**0.6%**	**-3.9%**	**10.2%**	**3.2%**	**1.4%**	**-1.3%**

Table 7
Average Relative Effective Spread by Market Capitalization Quintile (in Basis Points)

Bench-mark Quote Lag	Market Cap Quintile	True Side From Glue				LR Side		EMO Side	
		Order Submission Time Quote		Trade Time Quote					
		NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq
0	**1**	49.4	52.5	46.5	49.8	51.3	49.8	48.6	48.4
0	**2**	16.5	34.1	13.5	30.7	15.7	31.0	14.4	29.5
0	**3**	13.1	19.4	11.4	17.2	12.6	17.4	11.2	16.5
0	**4**	12.5	14.4	11.4	12.3	12.4	12.4	10.9	11.5
0	**5**	9.0	8.4	7.8	7.2	9.3	7.5	8.3	7.1
0	**Average**	**20.1**	**25.7**	**18.1**	**23.4**	**20.3**	**23.6**	**18.7**	**22.6**
2	**1**			46.6	52.9	52.0	52.9	49.4	51.4
2	**2**			14.7	33.5	17.0	33.9	15.7	32.3
2	**3**			12.6	19.4	13.9	19.9	12.4	19.0
2	**4**			12.0	13.8	13.3	14.1	11.6	13.1
2	**5**			8.3	7.8	10.0	8.5	8.9	7.9
2	**Average**			**18.8**	**25.5**	**21.2**	**25.9**	**19.6**	**24.8**
5	**1**			48.9	52.2	54.2	52.2	51.6	50.7
5	**2**			15.6	33.3	17.3	33.6	16.2	32.1
5	**3**			12.7	19.1	13.8	19.2	12.5	18.8
5	**4**			12.3	13.5	13.5	14.1	11.9	13.1
5	**5**			8.6	8.2	10.1	8.8	9.3	8.3
5	**Average**			**19.6**	**25.3**	**21.8**	**25.6**	**20.3**	**24.6**
10	**1**			49.2	52.2	53.9	52.2	51.9	50.7
10	**2**			16.0	33.0	18.1	33.2	16.8	31.7
10	**3**			13.3	18.3	14.4	18.4	13.2	18.2
10	**4**			12.5	13.5	13.5	14.4	12.2	13.4
10	**5**			8.8	8.4	10.4	9.0	9.5	8.5
10	**Average**			**20.0**	**25.1**	**22.0**	**25.4**	**20.7**	**24.5**

Table 8
Decomposition of Bias in Relative Effective Spread Estimates

Quote Lag	Market Cap Quintile	Midpoint Change Bias		Actual Bias				Algorithm Bias			
				LR		EMO		LR		EMO	
		NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq	NYSE	Nasdaq
0	**1**	-5.8%	-5.2%	3.9%	-5.2%	-1.6%	-7.7%	10.3%	0.0%	4.5%	-2.7%
0	**2**	-18.0%	-9.8%	-5.0%	-9.0%	-12.7%	-13.3%	15.9%	0.9%	6.5%	-3.9%
0	**3**	-12.7%	-11.3%	-3.2%	-9.9%	-14.3%	-14.9%	11.0%	1.5%	-1.7%	-4.0%
0	**4**	-8.6%	-15.0%	-0.6%	-14.3%	-13.1%	-20.3%	8.8%	0.8%	-4.9%	-6.3%
0	**5**	-13.0%	-13.8%	3.7%	-11.0%	-7.1%	-15.6%	19.2%	3.3%	6.7%	-2.0%
0	**Average**	**-11.6%**	**-11.0%**	**-0.2%**	**-9.9%**	**-9.8%**	**-14.3%**	**13.0%**	**1.3%**	**2.2%**	**-3.8%**
2	**1**	-5.5%	0.9%	5.3%	0.9%	0.0%	-2.0%	11.5%	0.0%	5.9%	-2.9%
2	**2**	-10.6%	-1.5%	3.4%	-0.5%	-4.6%	-5.1%	15.7%	1.0%	6.8%	-3.6%
2	**3**	-3.8%	0.1%	6.5%	2.6%	-5.3%	-1.8%	10.7%	2.4%	-1.6%	-2.0%
2	**4**	-3.9%	-4.6%	6.1%	-2.0%	-7.1%	-8.9%	10.5%	2.7%	-3.3%	-4.5%
2	**5**	-7.6%	-6.5%	11.5%	0.8%	-0.6%	-5.7%	20.7%	7.9%	7.5%	0.9%
2	**Average**	**-6.3%**	**-2.3%**	**6.6%**	**0.3%**	**-3.5%**	**-4.7%**	**13.8%**	**2.8%**	**3.0%**	**-2.4%**
5	**1**	-1.0%	-0.6%	9.8%	-0.6%	4.5%	-3.4%	10.9%	0.0%	5.6%	-2.9%
5	**2**	-5.2%	-2.1%	5.0%	-1.3%	-1.9%	-5.7%	10.8%	0.9%	3.5%	-3.7%
5	**3**	-2.6%	-1.4%	5.9%	-0.6%	-4.0%	-2.7%	8.7%	0.8%	-1.4%	-1.3%
5	**4**	-1.7%	-6.7%	7.7%	-2.2%	-5.1%	-9.5%	9.6%	4.9%	-3.4%	-3.0%
5	**5**	-4.6%	-2.0%	12.8%	5.0%	3.0%	-0.8%	18.2%	7.1%	7.9%	1.3%
5	**Average**	**-3.0%**	**-2.6%**	**8.2%**	**0.1%**	**-0.7%**	**-4.4%**	**11.6%**	**2.7%**	**2.4%**	**-1.9%**
10	**1**	-0.4%	-0.5%	9.2%	-0.5%	5.1%	-3.3%	9.6%	0.0%	5.4%	-2.8%
10	**2**	-2.8%	-3.1%	9.8%	-2.5%	1.8%	-6.9%	13.0%	0.6%	4.7%	-3.9%
10	**3**	2.1%	-5.6%	10.2%	-5.1%	0.7%	-6.1%	7.9%	0.5%	-1.4%	-0.5%
10	**4**	0.0%	-6.6%	8.0%	0.0%	-2.1%	-6.9%	7.9%	7.0%	-2.2%	-0.3%
10	**5**	-1.9%	-0.3%	15.2%	6.8%	5.7%	1.0%	17.4%	7.2%	7.7%	1.3%
10	**Average**	**-0.6%**	**-3.2%**	**10.5%**	**-0.2%**	**2.2%**	**-4.4%**	**11.2%**	**3.1%**	**2.9%**	**-1.2%**