Option Trading Costs Are Lower Than You Think*

Dmitriy Muravyev
Boston College

Neil D. Pearson
University of Illinois at Urbana-Champaign

March 15, 2015

Abstract

Conventionally measured bid-ask spreads of liquid equity options are large. This presents a puzzle, which we resolve. At high frequency, changes in option prices can be predicted using recent changes in stock prices. A large proportion of option trades exploit this predictability to take liquidity at low cost, buying and selling immediately before option prices are expected to change. Conventional measures of effective spreads and price impact do not account for this execution timing but can be adjusted to do so. For the average trade, effective spreads that take account of trade timing ability are one-third smaller than the conventionally measured effective spreads; for trades that reflect execution timing, they are five times smaller. These findings have striking implications for the profitability of options trading strategies that involve taking liquidity. In addition, conventional measures of price impact overstate it by a factor of more than two. Our results also indicate that most option trades originate from investors who time executions, for example proprietary traders and institutional investors who have access to execution algorithms.

Keywords: Execution timing, trading costs, effective spread, liquidity, equity options, algorithmic trading

---

* We gratefully thank Nanex and Eric Hunsader for providing the trade and quote data for the options and their underlying stocks. Most of the results in this paper were included in a previous working paper that circulated under the title "Negative Externality of Algorithmic Trading: Evidence from the Option Market." We thank Nicholas Hershey, Richard Payne, Christina Scherrer, and seminar or conference participants at Boston College, the Commodity Futures Trading Commission, the University of Illinois at Urbana-Champaign, the University of Southern California, Investment Technology Group, the University of Connecticut, the Workshop on High-frequency and Algorithmic Trading at the City University of Hong Kong, the Fifth Risk Management Conference at Mont Tremblant, and the European Winter Finance Conference for their comments and suggestions.
E-mail addresses: muravyev@bc.edu (D. Muravyev), pearson2@illinois.edu (N. D. Pearson)

Option Trading Costs Are Lower Than You Think

March 15, 2015

Abstract

Conventionally measured bid-ask spreads of liquid equity options are large. This presents a puzzle, which we resolve. At high frequency, changes in option prices can be predicted using recent changes in stock prices. A large proportion of option trades exploit this predictability to take liquidity at low cost, buying and selling immediately before option prices are expected to change. Conventional measures of effective spreads and price impact do not account for this execution timing but can be adjusted to do so. For the average trade, effective spreads that take account of trade timing ability are one-third smaller than the conventionally measured effective spreads; for trades that reflect execution timing, they are five times smaller. These findings have striking implications for the profitability of options trading strategies that involve taking liquidity. In addition, conventional measures of price impact overstate it by a factor of more than two. Our results also indicate that most option trades originate from investors who time executions, for example proprietary traders and institutional investors who have access to execution algorithms.

Keywords: Execution timing, trading costs, effective spread, liquidity, equity options, algorithmic trading

1. Introduction

At first glance, option market bid-ask spreads are puzzling. In our data on some of the most liquid and actively traded options, quoted option bid-ask spreads average 8.1 cents per share. Spreads are even wider for options that are well in-the-money. Conventionally measured effective spreads, which reflect the fact that trades tend to occur when quoted spreads are narrow, average 6.2 cents per share. For comparison, the average option price in our sample is $1.70. To the extent that these quoted and effective spreads measure the costs of taking liquidity in the options market, the costs of taking liquidity are high. In contrast, the spreads on the options' underlying stocks average 1.4 cents per share, and often are only one cent per share.

Puzzlingly, quoted option bid-ask spreads did not change much during our sample period of April 2003 to October 2006 despite the more than doubling of option trading volume from 2003 to 2006, by which time option trading volume was 17% of stock trading volume.[1] The failure of quoted spreads to decline appears to be inconsistent with both theories in which reductions in trading costs increase trading volume and theories in which increases in trading volume lead to lower costs per unit. Our sample period is also the period during which algorithmic trading came to dominate the option markets, making the failure of quoted spreads to decline even more surprising. Who is paying these high quoted spreads?

Existing theories are also unable to explain why spreads of options on the same underlying stock increase with option moneyness and the high spreads of in-the-money options. This pattern cannot be explained by hedge rebalancing costs incurred by option market makers, because hedges of well in-the-money options rarely need to be rebalanced. Similarly, the pattern cannot be explained by market makers' costs of hedging gamma and vega risks, because well in-the-money options are not exposed to these risks. Moreover, through the put-call parity relation in-the-money calls (puts) have gamma and vega risks similar to those of their corresponding out-of-the-money puts

[1] Data from the Options Clearing Corporation indicates that equity (not including index) options on approximately 83 and 184 billion shares traded during 2003 and 2006, respectively (http://www.optionsclearing.com/webapps/historical-volume-query). During 2013 option trading volume was 27% of stock trading volume computed using CRSP data, which includes trading in non-optionable stocks

(calls), but much different spreads. The large differences between the spreads of options and their underlying stocks also cannot be explained by differences in the adverse selection component of the spread, unless informed traders are much more common in the options market than in the stock market and they choose to trade in-the-money options rather than at- or out-of-the-money options with more embedded leverage. The low bid-ask spreads in the stock market imply that option spreads also are too large to be explained by option market makers' costs of executing the initial delta hedge trades.

The existing literature has attempted to explain option quoted and effective spreads using proxies for initial delta hedging costs, hedge rebalancing costs, and asymmetric information, and achieved limited success (Jameson and Wilhelm 1992, George and Longstaff 1993, Cho and Engle 1999, De Fontnouvelle et al 2003, Kaul et al 2004, Engle and Neri 2010, and Goyenko et al 2014). Most of these papers study either S&P 100 options or small numbers of equity options using short data samples, and most do not use data from the current market environment following the Options Linkage and the widespread adoption of algorithmic trading.[2] This literature uses regression analyses to produce evidence that proxies for the initial delta hedging costs, hedge rebalancing costs, and asymmetric information are correlated with option quoted and effective spreads. As discussed above, such analyses cannot explain why the levels of spreads are so large. The only paper that uses a large dataset from the current market environment, Goyenko et al (2014), separately considers in-, at-, and out-of-the-money options and thus does not attempt to explain the cross-section of spreads.

We resolve the puzzle of high option spreads by showing that the cost of taking liquidity in the option market is much less than both the quoted spread and the conventionally measured effective spread. These measures do not account for the fact that many investors take liquidity by buying (selling) options at times when recent stock price changes and other high-frequency public information imply that expected changes in option prices over very short horizons are positive (negative). A large fraction of the options trades in our sample, about 40%, reflect such trade timing; during the last sample month the fraction was 54%. For trades that display high-frequency trade timing ability,

[2] Goyenko et al (2014) use a large recent sample, while Engle and Neri (2010) use data from "nine liquid tickers in the financial sector traded in four dates of 2007."

an effective spread measure that takes account of it is only about 1.3 cents per share, just 21% of the conventionally measured effective spread of 6.2 cents per share and 16% of the average quoted spread of 8.1 cents per share. By the last month of the sample period the effective spread measure that takes account of trade timing ability is only 1.1 cents per share. These findings have striking implications for the possible profitability of options trading strategies that involve taking liquidity.

Averaging over all trades, the effective spread measure that takes account of high-frequency trade timing ability is just 67% of the average conventionally measured effective spread and 53% of the average quoted spread. The new measure of effective spreads declined during the sample period, beginning at 5.5 cents per share and reaching 3.5 cents per share by the end of the sample period. This decline was primarily driven by the increase in the fraction of trades that exploit timing ability, which almost doubled from 27.5% to 54% of trades.

At most only a few retail investors have the resources and ability to time their option trades based on high-frequency changes in stock prices and other market information. Thus, our finding that 40% of option trades exploit ability to time executions indicates that a large proportion and perhaps most option trading is done by sophisticated proprietary traders or institutional investors who either possess execution algorithms or have access to brokerage firm execution algorithms, and also suggest that the recent growth in option volume was driven primarily by professional investors entering the market. The 40% estimate is a lower bound on the proportion of sophisticated investors' option trades because some sophisticated investors will sometimes trade options without timing executions. Retail and other investors who are not able to time executions will on average trade when the option price is expected to stay the same and their costs will equal the conventionally measured effective spreads. During the last month of our sample period the fraction of option trades exploiting execution timing had increased to 54%, indicating that more than half of option trades are due to sophisticated investors.

Our lower estimates of the costs of taking liquidity are driven by the fact that at high frequencies option price changes can be predicted using recent changes in underlying stock prices and other high-frequency public information. We document this, and also that many option trades exploit this predictability to trade at favorable times. Buy trades are executed when high-frequency public information indicates positive expected short-term changes in option prices, and sell trades are executed when such information indicates negative expected changes in option prices. An alternative interpretation is that option traders buy after the value of the option has increased to be close to the ask price but the ask has not yet been adjusted, and sell after the value of the option has declined to be close to the bid price but the bid has not yet been adjusted. We use two simple models to estimate the expected change in the option price over short horizons, which, combined with the current option price, provide estimates of the expected future option prices based on public information.

The effective spread is estimated from the difference between the transaction price and an estimate of the fair market value or "underlying true value" of the security.[3] Conventionally, the bid-ask midpoint is used as an estimate of the security value, and thus the effective spread is measured as the difference between the transaction price and the bid-ask midpoint. The justification for this conventional approach is the lack of a readily available better estimate of the security value; if a better estimate is available, researchers should use it. We take up the challenge of developing a better measure of the security value. In particular, we use an estimate of the expected future price based on past publicly available information. Because trades tend to occur at times when our estimates of expected future prices based on public information are systematically different from the bid-ask midpoints, our estimates of the effective spread differ from conventional measures.

The execution timing that we document also resolves the puzzle of why dollar spreads of in-the-money options are so much larger than those of at- or out-of-the-money options. Option spreads have to be wide, or else movements in stock prices would create arbitrage opportunities as option market makers get "picked off." They are wider for options with larger deltas because the prices of such options are more sensitive to stock price movements. The large bid-ask spreads limit traders' ability to exploit option price

[3] For example, the recent survey by Bessembinder and Venkataraman (2010) explains that the effective spread is computed using "an observable proxy for the true underlying value of [the] security." In measuring price impact Hasbrouk (1991) uses the concept of the "efficient price," which is defined as the expected future price as the forecast horizon becomes large computed from computed from a vector auto-regression.

predictability to develop profitable stand-alone trading strategies. The large spreads also mean that taking advantage of the short-term predictability is a key element in reducing trading costs. In contrast to the quoted and conventional effective spreads, our estimates of the costs of taking liquidity in the in-the-money options are not implausibly high but rather are consistent with the costs of executing the initial delta hedge trades.[4]

Execution timing also has implications for estimates of price impact. Conventional measures of price impact are estimated from the difference between the midpoint sometime after the trade and the prevailing midpoint at the time of the trade. We replace the midpoint at the time of the trade with the expected future midpoint, a better estimate of the underlying security value, and obtain estimates of price impact that are only about one-half as large as conventional measures. Our findings about effective spreads and price impact corroborate the growing concern that traditional microstructure measures do not properly capture execution costs and price impact in modern electronic markets (Holden and Jacobsen 2013). In addition to corroborating this concern, we provide measures that take account of the high-frequency predictability of prices.

The main results about the differences in the costs of taking liquidity are confirmed using an alternative method to classify trades into those likely to have been initiated by execution timers (algorithms) and non-timers (human traders) that is not based on a model of future option price changes. Human traders are more likely than execution algorithms to choose round numbers (divisible by ten) as their trade size. Indeed, humans psychologically prefer round numbers (Rosch 1975), while algorithms often compute trade sizes using mathematical formulas. Taking advantage of this, we use non-round and round trade sizes as proxies for trades initiated by algorithms and directly by humans, respectively.[5] Trades of non-round size identified as likely algorithmic

[4] In most cases, investors who wish to establish or close out option positions have little alternative to taking liquidity and execution timing. For each underlying stock, option trading is spread across up to several hundred option contracts, most of which trade infrequently. As a result, a customer limit order in a particular option is unlikely to execute within a reasonable amount of time, and option market-makers provide liquidity in most option transactions.

[5] While this alternative identification of algorithmic and human trades does not depend on the model we use to predict option price changes, it is consistent with it. Our predictive model indicates that trades of round size (e.g., 30 contracts) are more likely to display execution timing ability than trades of similar but non-round sizes (e.g., 29 or 31 contracts).

trades have substantially larger conventionally measured price impact than the round trades identified as likely non-algorithmic trades. However, our price impact measure that takes account of execution timing is similar for both round and non-round trades. This implies that the difference in conventionally measured price impacts can be explained by execution timing.

Our results are important for interpreting research that documents the performance of option trading strategies. Recent such studies include Goyal and Saretto (2009), Driessen, Maenhout, and Vilkov (2009), Bali and Murray (2013), Cao and Han (2013), Doran, Fodor, and Jiang (2013), Boyer and Vorkink (2014), and Muravyev (2014). The low transaction costs obtainable via execution timing likely make profitable some trading strategies that would otherwise not be profitable. For example, Goyal and Saretto (2009) report that their long-short decile straddle portfolio returns are reduced from 22.7% to 3.9% per month if they assume that options are traded at the quoted spread, while Driessen, Maenhout, and Vilkov (2009) find that the alpha of their trading strategy becomes insignificant when they assume that trades occur at quoted bid and offer prices. Cao and Han (2013) report results for widely varying assumptions about effective spreads, presumably because they have little information about the costs of taking liquidity in the option market. Our finding that costs of taking liquidity are much less than quoted spreads and conventionally measured effective spreads, and very much less for traders who effectively time executions, has implications for analyses like these. We also contribute to the literature on optimal trade execution (e.g., Almgren and Chriss 2001 and Bertsimas and Lo 1998) by showing one mechanism that can be used to reduce trading costs. The paper also provides a rare glimpse into how some execution algorithms can operate at high frequencies.

2. Data

The paper uses tick-level data for 39 stocks including 2 ETFs from the option and equity markets. The data are provided by Nanex, a firm specializing in delivering high-quality data feeds. The sample period includes 882 trading days from April 2003 through October 2006. The selected stocks had the largest option trading volume during March 2003, just prior to the beginning of the sample period. The data include trades and best quotes for both stocks and options from all exchanges which list them. Muravyev, Pearson, and Broussard (2013) describe the data in more detail.

Our sample period begins shortly after the introduction of the Options Linkage connected all U.S. option exchanges in January 2003 and forced exchanges to upgrade their infrastructure, and was the period during which algorithmic trading came to dominate the U.S. options markets. The period of changes in the competitive landscape during which new options exchanges entered the market and almost all options became multiply listed was completed prior to the beginning of our sample period, and the reduction in tick sizes to pennies occurred after the end of the sample period. Mild data filters are applied to the trade sample. We include options with between 5 and 700 calendar days before expiration. The first and last five minutes of trading are excluded to avoid the opening and closing rotations. Trades for which implied volatility or the expected option price cannot be computed are also excluded. After applying all filters, the final sample consists of 20.4 million option trades. The Nasdaq ETF QQQ has the largest number of trades (1.8 million before the ticker change and 1.9 million afterwards) while AOL has only 52 thousand trades.[6]

Summary statistics are reported in Table 1. An average trade has a price of 1.70 dollars and size of 30 contracts on hundred shares each. However, the trade size distribution is highly skewed with 50th and 75th percentiles of 10 and 20 contracts respectively; and 14% of trades have the smallest possible size of one contract. There are slightly more seller-initiated trades (54%) than buyer-initiated trades (46%), and considerably more call option trades (64%) than put option trades (36%).

The trade direction is determined by the quote rule. If a trade price is at the quote midpoint of the National Best Bid and Offer (NBBO), then the quote rule is applied to the best quotes of the reporting exchange. The method is quite reliable as 84% of trades occur at the NBBO prices. On average, three out of six option exchanges quote the best national price at the time of a trade.

[6] AOL dropped from the sample after changing its ticker in October 2003. Several other stocks also dropped from the sample due to ticker changes.

3. Execution Timing in the Option Market

At high frequency changes in estimates of option values based on recent changes in stock prices predict future option price changes. We first explain our estimates of option values, and then document that short term changes in option prices can be predicted based on recent changes in the prices of their underlying stocks. After that we show that investors exploit this predictability to time executions of option trades.

3.1 Estimates of option values implied by underlying stock prices

The underlying stock price can be transformed using the Black-Scholes-Merton (BSM) formula into its option price equivalent, which we call the implied option price. We compute the implied option price by combining the current stock price with implied volatility estimated from past stock and option bid-ask midpoints.

The method consists of two steps outlined in Eq. (1):

$$\hat{P}_t^{\text{BSM}}(K,T) = \text{BSM}_t(S_t, IV_{t-}, K, T), \quad IV_{t-} = \frac{1}{N}\sum_{i=1}^{N} IV_{t-i\Delta t}\ , \tag{1}$$

where $\hat{P}_t^{\text{BSM}}(K,T) = \text{BSM}_t(S_t, IV_{t-}, K, T)$ is the option price computed using the BSM formula, S_t is the underlying stock price at time t, IV_{t-} is the average implied volatility over the previous thirty minutes, K is the option strike price, and T is the time to expiration. Specifically, for each option we compute the implied volatility using stock and option bid-ask midpoints at two-minute frequency over the previous 30 minutes, i.e. a total of N = 15 estimates, and then average the 15 estimates.[7] In the second step, the current stock price is transformed into the implied option price using the past implied volatility and the same BSM formula as in the first step.[8] Muravyev, Pearson, and Broussard (2013) use a similar idea.

The method can be viewed as a non-linear regression between the option and stock prices with one unknown parameter, the implied volatility. The regression is estimated on the recent price history and is then used to predict the option price corresponding to the current stock price. Thus, the method depends little on the particular

[7] The results depend little on the particular scheme used to compute implied volatility.

[8] As for the other parameters in the BSM formula, we assume no dividends and set the risk-free rate equal to 60-day LIBOR. Time to expiration is measured using calendar time. The results change little if we use a stock price with one second lag to allow for possible latency between the markets.

option pricing model and its assumptions. However, it does require two assumptions. First, it assumes that implied volatility changes more slowly than the underlying stock price during a trading day. Indeed, after adjusting for market microstructure effects, implied volatility usually changes slowly and smoothly intraday. Second, the implied option price should equal on average to the option quote midpoint during the estimation window (30 minutes), which is equivalent to assuming that the option quote midpoint is on average an unbiased estimate of the option fair market value.

3.2 Short-term option price predictability

The implied option price is a good predictor of the change in the option price over the next few minutes. We show this using a simple univariate regression of the change in the quote midpoint over a horizon of length τ on the difference between the implied option price and the quote midpoint,

$$P_{t+\tau} - P_t = \alpha_0 + \alpha_1(\hat{P}_t^{BSM} - P_t) + \varepsilon_t, \tag{2}$$

and then extend the model to include other predictors in Eq. (3) below. The model in Eq. (2) is estimated on regularly spaced five-second intervals over each trading day, pooling together all options on a given stock. We then average the coefficient estimates across all trading days for each stock, and compute *t*-statistics for the average coefficient estimates. Table 2 reports the results on a stock-by-stock basis.

The results in Table 2 show that changes in option quote midpoints over each of three time horizons (τ = 1 minute, 10 minutes, and 1 hour) are predicted by the difference between the implied option price and the option quote midpoint. The implied price explains a large portion of the short-term variation in option prices, with an average R^2 of 22% for the one-minute horizon.[9] The coefficients range from 0.34 to 0.69, with an average of 0.54. That is, in just one minute, the option price moves more than half the distance required to converge to the implied price. As expected, the average value of the regression coefficient is larger over the 10-minute and one-hour horizons.

Although this simple model works well, traders who time executions may use additional information to predict option price changes. We find that although other

---

[9] As expected, R^2 decreases with the time horizon from 10% for the ten-minute horizon to 3% for the one-hour horizon. Even the latter number is large for regressions that predict price changes or returns.

variables somewhat improve the forecasts, the implied price remains the most important predictor. The model

$$P_{t+\tau} - P_t = \alpha_0 + \alpha_1(\hat{P}_t^{BSM} - P_t) + \alpha_2(\hat{P}_t^{BBO} - P_t) + \alpha_3\#\mathrm{ExchBid}_t + \alpha_4\#\mathrm{ExchAsk}_t + \\ + \sum_{i=1}^{12}\alpha_{i+4}(\Delta_{t-i\Delta t} dS_{t-i\Delta t}) + \sum_{j=1}^{12}\alpha_{j+16} dP_{t-j\Delta t} + \varepsilon_t \tag{3}$$

extends the simple model in Eq. (2) and includes information about the limit order book and short-term option and stock price dynamics. The state of the limit order book is represented by the difference between the average quote midpoint across all exchanges (BBO average) and the NBBO quote midpoint, $\hat{P}_t^{BBO} - P_t$. We also include the numbers of exchanges at the best ask and bid prices with the idea, that if only a single exchange quotes the best ask (bid) price, it is likely to increase (decrease) soon. Price changes over the previous minute are represented by option and delta-adjusted stock price changes over the 12 most recent five-second periods. The regression is estimated separately for each stock and six groups of options defined by absolute delta (cut-offs of 0.35 and 0.65) and time-to-expiration (cut-off of 60 days) on each day using regular five-second time snapshots.

Table 3 reports the average coefficient estimates across all stocks for regressions estimated using ten minute and one hour time horizons. All of the average coefficient estimates are highly significant and have the expected signs. Changes in the option quote midpoint are highly predictable with R^2's ranging from 9% to 17% across the six groups at the 10-minute horizon and 8% to 18% at the one-hour horizon. The difference between the BSM implied price and the NBBO midpoint, $\hat{P}_t^{BSM} - P_t$, is the most important variable, consistent with the results in Table 2 showing that this variable alone is able to predict changes in option prices. The difference between the average BBO across exchanges and the NBBO, $\hat{P}_t^{BBO} - P_t$, is the second most important variable. It is highly correlated with the difference $\hat{P}_t^{BSM} - P_t$ but also provides some independent information. Consistent with Muravyev et al. (2013), the option market lags slightly behind the underlying stock, and the option midpoint is mean-reverting perhaps because of aggressive limit orders. The role of short-term price swings diminishes as time horizon increases.

We rely primarily on the extended model in Eq. (3) because option traders who time executions are likely to have access to other information in addition to the stock price. However, the results for the round-sized trades in Section 6 suggest that many investors rely only on the implied option price in timing their trade executions.

3.3 Execution timing when option prices are predictable

If the option price changes predictably and tends to move toward the implied option price, then the difference between the implied option price and current quote midpoint signals the best time to execute a trade. Specifically, investors who desire to buy options should execute purchases when the implied option price (i.e., the estimate of option value) approaches the ask price. If they do this then the difference between the transaction price they pay, the ask price, and the option value will be small. Similarly, investors who desire to sell options should execute sales when the implied option price approaches the bid price, because doing so will make the difference between the value and the trade price they receive small. We call this strategy execution timing.

This intuition can be generalized from the implied option price to the general predictive model of option price changes. The expected future price from a regression aggregates more information and thus is a better estimate of the underlying option value than the implied option price. Specifically, the model in Eq. (3) is estimated on past data, and then the coefficient estimates are multiplied by current values of the covariates to produce the predicted option price at a given horizon (e.g., one hour). This predicted price can be used in the same way as the implied option price: if the price is expected to increase (decrease), then it is good time to buy (sell).

Figure 1 is a stylized illustration of execution timing. The figure shows option prices (vertical axis) evolving in time (horizontal axis). The expected future midpoint (i.e., the option value) is shown in green. It evolves over time as the stock price, which is not shown, also evolves. We assume that the current quote midpoint (grey) eventually converges to the future expected midpoint (green) implied by the current price of the underlying. Execution timers will wait until the expected quote midpoint approaches the bid price (dark blue) to execute their sell trades, indicated by solid blue arrows. In this illustration investors timed their sales well because if they had waited longer the bid price they receive would have decreased.

This example also illustrates why conventional measures of the bid-ask spread and price impact that use the current quote midpoint overestimate trading costs. Because the conventional measure of the effective spread uses the quote midpoint, it assumes that the sell trades in the figure incur the same costs as hypothetical buy trades executed at the same time. However, such buy trades are a poor trading decision because the investor could have waited for the expected decrease in price to take place and then buy at a better price. That is, the conventional measure of the effective spread fails to account for price predictability. More specifically, when the sell trades occur the current quote midpoint is above its expected future value, which is the estimate of option value. Using the higher current quote midpoint as a proxy for the option value overstates the effective spread and price impact of the trade. The quote midpoint is on average significantly higher than its expected value at the time of sell trades, and price will decrease even if no sell trades occur.

Figure 2 uses the data to show that investors do in fact actively engage in executing timing: they buy right before the price is expected to increase and sell before it is about to decrease. The red line shows the predicted change in option price based on public information immediately before a trade computed using Eq. (3) with a 10-minute horizon. The predicted change is plotted as a function of the signed trade size, in dollars. For positive signed trade sizes the option price is expected to increase by approximately one cent, ranging from slightly below one cent for small trades to slightly above one cent for large trades. Following negative signed trades the expected option price change ranges from -0.5 cents to almost -1.5 cents, with the price change being about -1 cent for much of the range of trade sizes. The blue line shows the option price changes during the 10 minutes following a set of simulated trades for the same option and date at random times that do not overlap with the 10 minutes periods following the time of an actual trade. As expected, the average option price changes following these trades are close to zero. The difference between the red and blue lines is due to execution timing.

The green line shows the change in the option price midpoint from the time of a trade until 10 minutes after the trade. The predicted price change conditional on a trade (the red line) has the same sign as the actual change (the green line) but is smaller because not all traders time executions and because option prices change due to inventory and adverse-selection impacts as discussed further in Section 4.2.

3.4 What fraction of trades reflects execution timing?

The difference between the expected future price from the regression model $\hat{P}_{t+\tau,i}$ and the current quoted price $P_{t,i}$, computed for each trade and adjusted for the trade direction, is a measure of execution timing (*ET*). We normalize the difference by the effective bid-ask half-spread at the time of a trade to express the benefits of the execution timing as a percentage of one measure of trading costs. The following equation summarizes the definition:

$$ET_i = \frac{(\hat{P}_{t+\tau,i} - P_{t,i})\, I_{\text{buy/sell}}}{\text{Effective Bid Ask Spread}_{t,i} / 2} \quad , \tag{4}$$

where $I_{\text{buy/sell}} = 1$ if the *i*th trade is buyer-initiated and -1 if it is seller-initiated. The larger the execution timing measure is for a trade, the more likely it is initiated by a trader who times executions.

This measure provides a lower bound estimate for the extent of execution timing. Although our regression model captures the first-order variation in the expected price changes, some investors may develop a better predictive model. In this case, a better model will find opportunities to trade at low costs that a simpler model will miss. That is why, similarly to the quote midpoint versus the implied price case, the expected price for a better model will be systematically above the one from a simpler model for buyer-initiated trades.

Using this measure, we can estimate the share of trades initiated by execution timers. While execution timers submit their buy (sell) trades when the expected change in price is positive (negative), for other traders who do not time executions the price is expected to stay the same on average. We use the idea that these other traders have *zero* execution timing to estimate the market share of execution timers. Indeed, assuming that execution timers do not trade when the expected change is negative, all the trades with negative timing are initiated by others. If the distribution for the expected price changes for other trades is symmetric around zero, then the distribution can be reconstructed as a mirror reflection of the negative part which is observed. Thus, the total number of trades not reflecting execution timing is simply twice the number of trades with negative timing. Subtracting this quantity from total number of trades, we can obtain an estimate of the number of trades reflecting executing timing. Thus, market share of trades reflecting execution timing can be computed with the following formula:

$$\text{Share of Execution Timing} = 1 - \frac{2 \times \sum I_{ET_i \leq 1}}{\text{Total Number of Trades}}, \tag{5}$$

where I_A is an indicator function for the set A. Using this approach, we estimate that share of trades initiated by execution timers increased from 27.5% at the start of our data period in April 2003 to 54% in late 2006 as reported in Figure 3 and Table 4. These percentages can be considered lower bounds on the amount of algorithmic liquidity taking in the options market because not all algorithmic liquidity taking will reflect execution timing.

This result implies that by the end of our sample most of the trades are originated by sophisticated traders such as proprietary trading firms, hedge funds, and institutional investors who either have their own execution algorithms or have access to brokerage firm execution algorithms. Although the idea underlying execution timing is simple, only these investors have access to the technology needed to implement it. The estimated fraction of execution timers provides a lower bound for the fraction of such investors in the options market because the inherent limitations of the regression model cause it underestimate the prevalence of execution timing and because some sophisticated investors do not time executions of some or all of their option trades. Overall, these estimates are inconsistent with a common view that retail investors are responsible for most of option trading, and also suggest that the recent growth in option volume was driven primarily by professional investors entering the market.

4. Biases in Measures of Execution Costs and Price Impact

This section shows how conventional measures systematically overstate options trading costs and estimates of price impact when prices change predictably and options investors engage in execution timing

4.1 Effective spreads

Theoretical literature emphasizes that measures of trading costs should rely on the best estimates of securitie's fair market values. The expected future price from the regression model is the best linear estimate of this kind and thus should be used. However, the empirical literature overwhelmingly uses the current quote midpoint instead of its expected future value, implicitly assuming that price changes are unpredictable. Although the quote midpoint is a less precise estimate of the option value than the expected future quote midpoint, this does not create any bias if traders have no or limited ability to time executions. Indeed, the quote midpoint at a random moment is equal on average to its future value.

However, the situation differs in markets in which it is possible to time executions. Although the quote midpoint at a *random* moment is unbiased, it is biased *at the time of a trade* by a trader who times executions. As a result, conventional ex-ante measures of the bid-ask spread overstate trading costs. This bias can be corrected by replacing the quote midpoint with its expected future value from a regression model, i.e. the usual effective spread measure $2I_{B/S}(TP_t - P_t)$ is replaced by the adjusted effective spread $2I_{B/S}(TP_t - \hat{P}_{t+\tau})$, where TP_t is the trade price at time t and $\hat{P}_{t+\tau}$ is the estimate of the midpoint at time $t + \tau$.

Consider an example. Suppose that a call option is trading at 1.0/1.1 dollars bid/offer, and the quote midpoint is expected to increase by 1 cent in the next minute from 1.05 to 1.06 dollars. An investor wants to buy at the ask price because the price is about to increase. The conventional measure of the effective half-spread for this trade is 5 cents (1.10 – 1.05), while the actual cost as measured by the adjusted effective half-spread is only 4 cents (1.10 – 1.06).

Table 5 reports different trade-weighted measures of the spreads for each stock in our sample, along with the averages across stocks (at the top of the table).[10] The column headed "Avg. Quoted" reports the average daily quoted spreads based on the NBBO. This average is computed by assigning to each option trade the average quoted spread for the day, where the quoted spreads are computed separately for each option from one-

[10] The overall average spread measures in this table differ from those in Table 4 due to a difference in the way the averages are computed. The average spreads in the first row of Table 5 are averages across stocks, where each stock is weighted equally. The average in Table 4 are across months, and do not include the stocks that dropped from the sample in the months after the stocks dropped.

second snapshots on each day. These reflect trading costs for an investor who trades at random times.[11] The overall average of 8.4 cents in the first row indicates that such an investor pays 8.4 cents for a round-trip trade or 5 percent of an average option price of 1.70 dollars. Investors can reduce the costs by trading when the quoted spread is narrower, and trades tend to occur when quoted spreads are less than average. The column headed “Average Quoted at Time of Trade” shows that the average quoted spread at the time of a trade is 6.6 cents per share. The average effective spread in the next column is the doubled difference between the trade price and the quote midpoint. It is just a bit smaller, 6.4 cents per share, reflecting the fact that occasionally trades occur inside the quoted spread, though this is not frequent as more than 80% of option trades are executed at the NBBO quotes. The next column shows the average realized spread.

The next two columns headed “Adjusted, BSM” and “Adjusted, Regression” show adjusted effective bid-ask spreads based on the simple regression model using only the difference between the BSM implied option price and the quote midpoint (Eq. 2) and the more general regression model (Eq. 3), respectively. The adjusted-effective bid-ask spread is twice the difference between transaction price and the expected price implied by the predictive model. The overall average estimate of the adjusted effective spread based on the regression model in the column “Adjusted, Regression” is 4.5 cents, which is 30% less than the conventionally measured effective spread and 46% less than the average quoted spread. The spread computed using the BSM implied option price is even smaller, 4.2 cents, which is only half of the quoted spread. The execution timing affects not only the level of trading costs but also the relative stock ranking. For example, Pfizer and QLogic have the same adjusted-effective spreads of 4.3 cents, but very different quoted spreads of 7 and 9.8 cents.

These results indicate that option trading costs are much lower than indicated by either quoted or conventionally measure effective spreads. Execution timing is essential for trade execution and significantly reduces trading costs. Further, these results are based on all trades, including trades executed by investors who do not engage in

[11] The literature mostly uses the end-of-the-day bid-ask spread from OptionMetrics, which is a special case of the average quoted spread with only one observation per day.

execution timing. The execution costs of traders who are able to time executions are lower still.

The role of the execution timing increased substantially during the sample period as execution algorithms improved and their share increased. Figure 4 shows that the adjusted effective spread decreases from 6.5 cents to 3.5 cents while the average quoted spread remains approximately unchanged at about 8 cents, and the effective spread modestly decreases from 7.5 to 6 cents. Thus, the adjusted-effective spread decreases by almost half while the conventional spreads change little. Several time-series properties of the average spreads are worth noting. Although trading costs for any particular stock are quite volatile, their market average fluctuates in a narrow range. Thus, trading cost volatility seems to be diversifiable at least during normal times. The average spreads follow a long-term trend and display little volatility clustering.

The adjusted effective spread is similar to the realized spread, with the difference being that the realized spread uses the actual post-trade midpoint $P_{t+\tau}$rather than the estimate $\hat{P}_{t+\tau}$. However, it has several advantages over the realized bid-ask spread. First, the adjusted spread is an ex-ante measure that can be used for trade execution. Second, the realized spread reflects not only trading costs but also the information and inventory impacts of a trade making it hard to disentangle these effects, while the adjusted spread only measures trading costs. Finally, the realized spread provides a more volatile estimate of trading costs because future price is more volatile than its forecast.

4.2 Measures of Price Impact

Prices respond to trades swiftly and by large amounts. The two columns of Table 6 headed “Observed Price Impact, Cents” show the conventional price impact measures $I_{B/S}(P_{t+\tau} - P_t)$ for the various stocks in our sample using horizons τ of one and ten minutes. The averages for the two different horizons across stocks are shown in the first row of the table.   These results show that, on average, the quote midpoint moves by 1.13 and 1.34 cents in the first one and ten minutes after a trade, which is large relative to the $1.70 average option price.

However, conventional measures of price impact significantly overestimate the causal effect of trades on prices, for the same reason that conventionally measured effective spreads do. Eq. (6) decomposes the observed price impact into the correctly measured price impact of a trade, and the expected change in the quote midpoint if no trade occurred:[12]

$$\underbrace{P_{t+\tau} - P_t}_{\text{Observed Price Impact}} = \underbrace{P_{t+\tau} - \hat{P}_t^{t+\tau}}_{\text{Price Impact}} + \underbrace{\hat{P}_t^{t+\tau} - P_t}_{\text{Expected Price Change}} \tag{6}$$

Further decomposing the price impact into components due to asymmetric information and inventory risk, we obtain

$$\underbrace{P_{t+\tau} - P_t}_{\text{Observed Price Impact}} = \underbrace{\alpha_{AI}(P_{t+\tau} - \hat{P}_t^{t+\tau})}_{\text{Asymmetric Information}} + \underbrace{\alpha_{IR}(P_{t+\tau} - \hat{P}_t^{t+\tau})}_{\text{Inventory Risk}} + \underbrace{\hat{P}_t^{t+\tau} - P_t}_{\text{Expected Change}}, \tag{7}$$

where $\alpha_{AI} + \alpha_{IR} = 1$

In the options market, the expected price change is of roughly the same magnitude as the correctly measured price impact of a trade if one uses a short horizon of one minute and is larger than the correctly measured price impact if one uses a horizon of ten minutes. Specifically, in Table 6 the pair of columns headed “Observed Price Impact” report estimates of conventionally measured price impacts for horizons of one and ten minutes for each of the stocks in the sample, as well as the averages across stocks. The pair of columns headed “Expected Price Change” and “Ajusted Price Impact” report estimates of the expected price changes and correctly measured adjusted price impacts for each of the stocks in the sample for the same horizons. For the one minute horizon the average expected price change is 0.47 cents, which is 42% of the conventionally measured price impact of 1.13 cents, and the correctly measured adjusted price impact is only 0.66 cents, or 58% of the conventionally measured price impact. For the 10 minute horizon, the average expected price change is 0.82 cents, which is 61% of the conventionally measured price impact of 1.13 cents, and the correctly measured adjusted price impact is only 0.52 cents, or 39% of the conventionally measured price impact. (In untabulated results using only the BSM implied option price with Eq. (2) we find that even without a trade the quote midpoint would move by 1.08 cents over the 10-minute horizon, which is 81% of the observed price impact of 1.34 cents.) Thus, although it is

[12] Price impact is adjusted for the trade direction everywhere in the paper. The expected part is estimated from a predictive model but with a smaller time horizon (for example, ten minutes) than commonly used in price impact measures.

tempting to attribute the large conventionally measured price impact to informed trading, in fact, over a horizon of 10 minutes the expected price change constitutes the majority of the conventionally measured price impact. The additional decomposition in Eq. (7) makes clear that the expected price change is large relative to the effects of both asymmetric information and inventory risk.

How price impact changes with trade size is of particular interest for identifying informed trading. Figure 6 shows the conventionally measured price impact exceeds one cent even for small trades. It is increasing for small trades and is almost flat (at approximately two cents) for trades of more than thirty contracts. But the most pronounced pattern, which is discussed in Section 6, is that trades of round (divisible by 10) sizes have significantly lower (by half a penny) price impact than non-round trades. To identify the causal impact of trades on prices the observed price impact should be adjusted by subtracting the expected price changes. These adjusted price impacts based on both the regression and BSM models are shown in Figure 7. Using the regression model, the adjusted price impacts are much smaller and now increase steeply with trade size. The BSM-adjusted price impacts are even smaller, and have many desired properties. They start almost from zero as trades of one contract have an adjusted price impact of only 0.07 cents. The BSM-adjusted price impacts increase monotonically to about 0.6 cents. Also, as discussed in Section 5, the differences between the price impacts of round and non-round trades disappear.

We use regression analyses to study how the observed and adjusted price impacts depend on trade characteristics. The first three columns in Table 7 shows the results from three different regressions of the observed price impact over a ten minute horizon following the trade on a number of variables, including the absolute value of the option delta ($|\Delta|$), the square root of the time to option expiration ($\sqrt{T-t}$), a dummy variable taking the value of one if the option traded is a call (I_{Call}), a time trend measured in years (TimeTrend), the option bid-ask spread measured in cents (Bid-Ask), a dummy variable taking the value of one if the trade is a purchase (I_{Buy}), the square root of the trade size measured in contracts ($\sqrt{\text{Size}}$), a dummy variable taking the value one if the trade size is one contract ($I_{\text{Size = 1}}$), the number of option exchanges at the NBBO on the side of the market where the trade occurred (#ExchAtNBBO), a dummy variable taking the value one if there is only one exchange at the NBBO on the side of the market where the trade occurred ($I_{\#\text{Exch} = 1}$). Two of the specifications also include the predicted price change from the regression model, $\Delta\hat{P}_t^{t+\tau}$, and one of the specifications includes dummy variables taking the value one if the expected quote change based on the regression model is between zero and two cents, two and five cents, and greater than five cents ($\Delta\hat{P}_t^{t+\tau} I_{0<x<2}$, $\Delta\hat{P}_t^{t+\tau} I_{2<x<5}$, and $\Delta\hat{P}_t^{t+T} I_{5<x}$, respectively). In the first specification that does not include any of the expected price change variables $\Delta\hat{P}_t^{t+\tau}$, $\Delta\hat{P}_t^{t+\tau} I_{0<x<2}$, $\Delta\hat{P}_t^{t+\tau} I_{2<x<5}$, or $\Delta\hat{P}_t^{t+T} I_{5<x}$ the most important variables are the two variables #ExchAtNBBO and $I_{\#\text{Exch} = 1}$ describing the state of the options limit order book. The coefficients of –0.571 and 0.562 on these variables indicate that for buy (sell) trades, the price impact decreases (increases) by 0.571 cents with each additional exchange at the ask (bid) price and increases (decreases) by 0.562 cents if a single exchange quotes the best ask (bid) price. Turning to the other coefficient, the time trend is very strong: the observed price impact increases by 0.55 cents each year. Price impact is increasing in trade size, but the impact of a trade of 100 contracts is only $0.019 \times \sqrt{100} = 0.19$ cents. The coefficient for the level of the option price is small (0.158), validating our approach of measuring option price impact in dollar terms.

The second column reports the results of a specification that also includes $\Delta\hat{P}_t^{t+\tau}$, the predicted price change from the regression model. The estimated coefficient on this variable is large, 0.776, and highly significant (*t*-statistic = 66.92), confirming that much of the conventionally measured price impact is explained by the expected price change based on public information.[13] Including this variable in the regression has important impacts on many of the other regression coefficient; for example, the coefficients on the variables #ExchAtNBBO and $I_{\#\text{Exch} = 1}$ describing the state of the options limit order book change from –0.571 to –0.198 and from 0.562 to 0.194, respectively. The third column also includes the variables $\Delta\hat{P}_t^{t+\tau} I_{0<x<2}$, $\Delta\hat{P}_t^{t+\tau} I_{2<x<5}$, or $\Delta\hat{P}_t^{t+T} I_{5<x}$ that capture non-linearities in the relation between the observed price change and the regression-based

[13] In a univariate regression that includes only $\Delta\hat{P}_t^{t+\tau}$ on the right-hand side the estimated coefficient on this variable is 0.92.

forecast of the price change. The estimated coefficients on these variables are highly significant, and the coefficient on $\Delta\hat{P}_t^{t+\tau}$ is reduced to 0.584.

The fourth column reports results for a specification in which the left-hand side variable is the BSM-adjusted price impact $\Delta P_t^{t+\tau} - \Delta\hat{P}_t^{BSM}$. Strikingly, for the BSM-adjusted price impact, most of the coefficients on the independent variables become much smaller in magnitude and the R^2 drops to zero. For example, the coefficient on the number of exchanges at the NBBO, #ExchAtNBBO, decreases from −0.572 in the first specification to −0.058, and the coefficient on $I_{\#\text{Exch} = 1}$ decreases from 0.562 to −0.006 and becomes insignificant. These observations together with the analysis of round and non-round trades in Section 6 are consistent with the hypothesis that option market-makers and algorithmic traders time executions using the BSM model or a similar approach.

4.3 Effective spreads of traders who do and do not time executions

We now turn to estimating the spreads of liquidity-taking trades that do and do not reflect execution timing and examining how they changed during our sample period.

Our approach to estimating the spreads of liquidity-taking trades that do and do not reflect execution timing is based on the tautological assumptions that (a) liquidity-taking trades executed by traders who time executions have non-negative execution timing, and (b) liquidity-taking trades by traders who do not time executions have no execution timing ability. We used these assumptions above to identify the shares of liquidity-taking trades by execution timers and non-timers. In particular, each trade is assigned a probability of being initiated by an execution timer. Our approach consists of two steps. First, we recover the empirical distribution of execution timing for the liquidity-taking trades that do not display execution timing. The second assumption (b) implies that the average execution timing should be zero or close to zero and its distribution should be symmetric around zero in a large sample of trades by traders who do not time executions Thus, only trades by traders who do not time executions can have non-positive execution timing. We use this subsample of trades with non-positive timing to infer the left half of the probability distribution for the execution timing of non-algorithmic traders. Because the distribution is symmetric, estimating its left side is equivalent to recovering the entire distribution. In the second step, we combine the execution timing distributions for all trades $f_{\text{All}}(ET)$ and the trades of non-timing investors $f_{\text{Non-timer}}(ET)$ to infer the execution timing distribution for the traders who do time executions $f_{\text{Timer}}(ET)$. Finally, the probability that trade i with execution timing ET_i is initiated by a non-timer is

$$P(i = \text{Non-timer}) = \frac{f_{\text{Non-timer}}(ET_i)}{f_{\text{Non-timer}}(ET_i) + f_{\text{Timer}}(ET_i)}, \tag{8}$$

where $f(\bullet)$ is the empirical density function. According to our assumptions, all trades with non-positive execution timing are initiated by non-timers. The likelihood that a trade is initiated by a trader who does not time executions decreases with the level of execution timing and is close to zero for trades with large positive execution timing. Appendix A explains the details of our approach.

Using this approach, we estimate the adjusted and conventional effective spreads for traders who do and do not time executions, and also the overall average spread, for each 20-day period that falls within our total sample period. Panel A of Table 4 reports adjusted effective spreads for the first and last 20-day periods ("months") of our sample period, as well as the overall averages in the row labelled "All Months." Panel A also reports the percentage of trades initiated by execution timers in the rightmost column.

As expected, traders who time executions pay significantly lower adjusted spreads than non-timers. The overall average adjusted effective spread paid by execution timers was 1.3 cents, compared to an overall average of 6.2 cents for traders who do not time executions. The adjusted effective spreads they paid decreased during the sample period from 1.9 to 1.1 cents per share, or by 0.8 cents per share. The adjusted effective spreads paid by non-timers also declined, from 6.8 to 6.0 cents per share. This decline in adjusted effective spreads paid by non-timers (and the similar reductions in Panel B) suggests that the reduction in effective spreads paid by execution timers was not driven by improvements in their ability to time executions, but rather by a decline in the average level of quoted spreads. The adjusted effective spread for the combined population of timers and non-timers in the column headed "All" declined from 5.5 to 3.5 cents per share, primarily driven by the increase in the proportion of trades that reflected execution timing from 27.5% to 54.0% of trades.

The striking results in Panel A are that the adjusted effective spreads paid by traders able to time executions averaged only 1.3 cents per share. Panel B of the same table shows that the conventional effective spread for the same set of traders was 6.2 cents, and thus overstates their trading costs by a factor of almost five! By the end of our sample period the adjusted and conventional spreads were 1.1 and 6.0 cents per share, respectively. The conventional measure of effective spreads gives a strikingly misleading answer about the option trading costs of professional traders who are able to time executions.

5. The pattern of option bid-ask spreads

Execution timing explains the main stylized fact about option bid-ask spreads, namely why dollar spreads increase so much in option moneyness. Figure 5 shows that for OTM options the average quoted spread is below 7 cents, while for ITM options the average quoted spread is 11 cents. By contrast, the adjusted effective spread is much flatter and increases from 4 to only 6 cents. For large trades, the relationship becomes completely flat with a 5 cent spread for ITM options. The residual 1.5 cent difference between ITM and OTM options is likely due to the initial hedging costs. Cho and Engle (1999) argue that option market makers immediately delta hedge after each trade and thus pay the spread in the underlying stock, which in our sample is 1.4 cents. To provide context for this result, Petrella (2006) shows that the initial-hedging theory of Cho and Engle (1999) cannot explain the difference between the spreads of ITM and OTM options because equity spreads are small. Explanations in terms of the costs of hedge rebalancing (Kaul, Nimalendran, and Zhang, 2004) also cannot explain the high spreads of in-the-money options because these options are precisely the ones for which hedge rebalancing is not needed. Nor can adverse selection about volatility information because the values of well in-the-money options are not exposed to the risk of changes in volatility. The execution timing ability of algorithmic liquidity takers is the only story that explains why option bid-ask spreads increase in the absolute value of option delta.

Finally, we also study how the execution timing bias measured by Eq. (4) depends on trade characteristics. Table 8 presents the results of several different specifications in which the execution timing bias is regressed on variables that describe the option trade and the state of the limit order book at the time of the trade. (These are the variables used in the regressions reported in Table 7, with the exception of the predicted price change and the other variables constructed from the predicted price change.) The timing bias is increasing in absolute value of the option delta because ITM options are more sensitive to changes in the underlying price and thus more exposed to execution timing. The average timing bias is 0.38 (or 38%), and the change from OTM (delta=25) to ITM (delta=75) increases the bias by 10%. Option market makers are aware of this effect as absolute delta becomes insignificant after controlling for the number of exchanges quoting the best price. The timing bias increases by 16% each year reflecting the growth in algorithmic trading. As expected, the number of exchanges quoting the best price in the direction of a trade is a significant determinant of the bias, with each additional exchange reduces the bias by 20%. In a special case of only one exchange quoting the best price the execution timing bias is larger by an additional 15%. The effects of the other explanatory variables are economically small.

6. Trades of round and non-round size

We compare the price impacts of trades of similar round and non-round sizes, for example trades of 30 contracts versus trades of 29 or 31 contracts. A trade is defined to be round if the trade size (in contracts) is greater than 15 and divisible by 10. Non-round trades are more likely to originate from execution algorithms and thus more likely to display execution timing. Human traders prefer round numbers and are more likely to acquire the target position in a single trade, implying that their trades will frequently be of round size. On the other hand, use of algorithms is likely to be correlated with the use of more sophisticated strategies in which trade size is determined by computations that do not result in round trade sizes. Also, execution algorithms sometimes split the target trade size into multiple child orders to minimize price impact and to take advantage of the quantities available at the best bid or offer quotes.

As discussed in Section 4.2, the conventionally measured price impact has three components: inventory risk, asymmetric information, and the expected price change due to execution timing (see Eq. 7). Round and non-round trades of similar size (for example, 30 vs. 29 or 31) carry similar amounts of inventory risk. The hypothesis that execution timing biases conventional measures of trading costs implies that non-round trades should have greater conventionally measured price impact as compared to similarly sized round trades, unless round trades reflect more information. Failure to find this would be evidence inconsistent with the claim that execution timing has important impacts on measures of price impact, again unless round trades contain more information. Thus, the comparison of round and non-round trades provides a test of the importance of execution timing that is not dependent on the BSM and regression models that we use to forecast option price changes. In addition, the adjusted price impacts computed using our models predicting price changes provide measures of the differences in information between non-round and round trades.

Figure 6 plots the average conventionally measured price impact for each trade size, along with the average expected price changes computed using the BSM and regression models.[14] Close examination of the line showing the average conventionally measured price impacts reveals that it has downward spikes at round trade sizes. The line also has smaller downward spikes for round-five trades, defined as those for which the trade size (in contracts) is greater than 15 and divisible by five. The downward spikes are approximately the same size for each round size, except that the downward spike for trades of 100 contracts is larger than the others.

The first column of results in Table 9 estimates the differences in the conventionally measured price impacts of round and round-five trades as compared to non-round trades by regressing the price impacts on a dummy variable that take the value of one for round trades and a second dummy variable for round-five trades that are not also round trades. The regression also includes as control variables several functions of trade size, measured in contracts, and the option price. The coefficient on the dummy variable for round trades is –0.411, indicating that the price impact of round trades is about 0.4 cents smaller that for non-round trades. This is a significant fraction of the average price impacts, which are 1.4 and 1.8 cents for round and non-round trades, respectively. Figure 6 indicates that the difference in the price impact of round and non-round trades is approximately the same for each round number, except that the difference

[14] Of trades larger than fifteen contracts, 60% have size divisible by ten which is six times larger than would be implied by a uniform probability distribution over trade sizes. It is difficult to explanation for why there are so many round-ten trades.

is larger for trades of 100 contracts. These results cannot be explained by an inventory effect, as round and non-round trades of similar sizes should have the same inventory risk, but are consistent with the hypothesis that execution timing is a large fraction of conventionally measured spreads. They are also consistent with the hypothesis that non-round trades contain more information about option values.

For the round-five trades the estimated coefficient is –0.199, indicating that the conventionally measured price impact of round-five trades is about 0.2 cents less than the non-round five trades. Perhaps coincidentally, the magnitude for round-five trades is half the magnitude for round-ten trades.

Examination of the other two lines in Figure 6 reveals that the expected price changes based on the BSM and regression models display downward spikes at the round and round-five trade sizes, the same locations as the downward spikes in the conventionally measure price impacts. That is, the variation in the conventionally measure price impact is matched by corresponding variation in the expected price changes based on public information computed from the BSM and regression models. From the figure, it appears that a large fraction of the variation in the conventionally measured price impact is explained by variation in the expected price changes computed using the two models.

Figure 7 pursues this issue by showing the adjusted price impacts based on the regression and BSM models. As in Section 4 above, in each case the adjusted price impact is the difference between the conventionally measured price impact and the model forecast of the price change at the time of the trade. The downward spikes at round trades in the adjusted price impacts based on the regression and BSM models are either smaller or much smaller than those in Figure 7. The second and third columns of Table 9 report the results of regressing the adjusted price impacts on the dummy variables for round trades and round-five trades that are not also round trades and the controls. The magnitude of the round trade effect in the adjusted price impacts based on the regression model decreases from –0.414 to –0.126 cents and the magnitude of the round-five effect in the regression-adjusted price impacts decreases from –0.199 to –0.045 cents. The results in the third column show that adjusting the price impacts using the BSM model is even more effective—the differences in the BSM-adjusted price impacts between round and non-round trades are small and not significantly different from zero. Strikingly, as shown in Figure 7 the BSM model eliminates the difference not only jointly but for almost any round size, with the larger round trades of size 90 and 100 being exceptions.

If non-round trades contain more private information than round and round-five trades, then the difference in the adjusted price impacts would remain large regardless of the execution timing model used to adjust the price impacts. On the other hand, if non-round trades do not contain more private information than round and round-five trades and one carries out the adjustment using the execution timing model used by algorithmic traders then the adjusted price impacts of round and non-round trades will be the same. Our finding that the differences between the price impacts of round and round-five trades as compared to non-round trades disappear after the price impacts are adjusted using the BSM model to account for the expected changes in price implies that non-round trades do not contain more private information than round and round-five trades. The larger observed price impact of non-round trades is entirely driven by their execution timing ability, and non-round trades contain the same amount of private information as round trades.

Our results also indicate that algorithmic traders who take liquidity and time executions use the BSM model, or something similar to it, for execution timing. The BSM-adjusted price impacts also have two other desirable properties that the regression-adjusted price impacts do not share. First, they start from almost zero for small trades and increases smoothly in size, as shown in Figure 7. Second, unlike other price impact measures, the BSM-adjusted price impacts do not depend on the number of NBBO exchanges and the time trend (last column of Table 7).

Because the round and non-round trades of similar sizes contain the same amount of private information, the estimates of the differences in price impacts provides a model-free lower bound for the difference in costs of taking liquidity between algorithmic and non-algorithmic traders. They prove a lower bound because algorithms submit both round and non-round trades, so that the round trades are not submitted exclusively by non-algorithmic traders. The large difference in price impacts eliminates a possible concern that our results are somehow mechanically driven by a particular model for the expected option price change. Overall, the round-trade analysis clearly shows that price impact measures should be adjusted to account for the expected changes in price to avoid spurious results about informed trading and inventory risk.

7. Conclusion

We show that execution timing is an important part of option trading. The conventional measure of the effective spread overestimates the average costs of taking liquidity by about 30% and the quoted spread overstates the cost of taking liquidity by a factor of almost two. For the traders who are able to time executions the costs of taking liquidity are even lower. By the end of our sample period the trading costs as measured by the adjusted effective spreads of the traders able to time executions were 1.1 cents per share, less than one seventh of the average quoted bid-ask spread and about one fifth of the conventionally measured effective spread of the same set of traders. The conventional measure is biased because it relies on the quote midpoint as a benchmark, whereas traders who are able to time executions tend to trade at times when the quote midpoint differs systematically from the option value and publicly available information indicates that option prices are expected to change. We show how to correct this bias by replacing the quote midpoint in these measures with the expected future value of the option based on publicly available information.

The conventional measure of price impact also uses the quote midpoint, and overstates price impact by a factor of about two.

Our finding that that by the end of our sample period 54% of option trades exploit ability to time executions indicates that most option trading is done by sophisticated proprietary traders or institutional investors who either possess execution algorithms or have access to brokerage firm execution algorithms, and also suggest that the recent growth in option volume was driven primarily by professional investors entering the market.

Our results about trading costs are important for studies that document option pricing anomalies and then examine their profitability after taking account of transactions costs. They also help explain why option trading is so popular despite wide quoted option spreads—for the majority of traders, the costs of taking liquidity are much less than quoted spreads. Finally, we show that execution timing explains the puzzle of why dollar spreads of in-the-money options are so much larger than those of at- or out-of-the-money options.

Our results also have policy implications for market design. In classic models of asymmetric information such as Copeland and Galai (1983), market-makers set the bid-ask spread so that losses from trading with against some traders are subsidized by gains from trading against others, including retail investors. The differences in execution timing ability imply that the realized spreads of the traders who are and are not able to time executions differ. This in turn implies that execution timing by some traders increases the trading costs of those who are not able to time executions. This observation immediately implies that the quoted bid-ask spread would be smaller without execution timing or, more generally, if the asymmetry between timing and non-timing traders is reduced. The asymmetry friction is driven not by private information, but by the ability to quickly process public information. One possible solution is to quote option prices in implied volatilities rather than dollar terms, as implied volatilities are less sensitive to changes in the underlying price than are option prices. A less radical solution would be to encourage exchanges to introduce limit orders linked to implied volatility.

Our approach for improving measures of trading costs by developing better estimates of value to use in place of the bid-ask midpoint is particularly well-suited to the option market because the prices of underlying stocks are key inputs to option prices, and at high frequency changes in stock prices predict changes in option prices. Nonetheless, it seems likely that the method can be extended to other markets in which there is value-relevant high-frequency information that is reflected in market prices with even a slight lag. For a stock, such value-relevant information might include, for example, returns of liquid stocks in the same industry or an industry exchange-traded fund.

Appendix A. Likelihood that a trade is initiated by an execution timer

As outlined in Sections 3.4 and 4.3, our approach to estimating the likelihood that a transaction is initiated by a trader who do not use execution timing, whom we call "non-timer" for brevity here, is based on the assumptions that (a) trades by execution timers have non-negative execution timing, and (b) trades by non-timers have no execution timing ability. The second assumption (b) implies that the average execution timing should be zero or close to zero and its distribution should be symmetric around zero in a large sample of trades by non-timers.

First, we recover the empirical probability distribution of execution timing for trades of non-timers. Based on the assumption all trades with non-positive execution timing are initiated by non-timers, we estimate the left half of the distribution from the subsample of trades with non-positive timing. Specifically, in each 20-day period, we first use the linear regression model of execution timing (Eq. 3) to estimate the execution timing of each trade. For each stock, we then sort the subsample of trades with non-positive execution timing estimates into 20 equally-sized portfolios P_{-20} to P_{-1}, where portfolio −20 consists of the trades with the worst (most negative) execution timing and portfolio −1 consists of the trades with zero execution timing. The boundaries $ET_{-20:0}$ (where $ET_0 = 0$) and the number of trade observations $N_{-20:-1}^{non-timer}$ in the portfolios describe the negative/left half of the distribution of execution timing of non-algorithmic traders. Because the distribution is assumed to be symmetric, the 20 portfolios for the positive/right half of the distribution $P_{1:20}$ will have the same number of observations (e.g., $N_{10}^{Non-timer} = N_{-10}^{Non-timer}$) and the same boundaries but with reversed sign (e.g., $ET_{10} = -ET_{-10}$). Thus, we have extracted the distribution for non-timers. The distribution of execution timing for timers can be extracted by simply subtracting the estimate for the number of non-timer trades from the actual number of trades in each portfolio $N_{10}^{Timer} = N_{10}^{Total} - N_{10}^{Non-Timer}$. By construction, the number will be zero for all portfolios with negative timing. Finally, the probability that a trade from i-th portfolio was initiated by non-timing investor is $(i = Non - Timer) = N_i^{Non-Timer}/N_i^{Total}$.

References

Almgren, Robert and Neil Chriss. 2001. “Optimal Execution of Portfolio Transactions.” *Journal of Risk* 3 (April), 5–39.

Bali, Turan G., and Scott Murray. “Does risk-neutral skewness predict the cross-section of equity option portfolio returns?” *Journal of Financial and Quantitative Analysis* 48.04 (2013): 1145-1171.

Bertsimas, Dimitras, and Andrew Lo 1998 “Optimal control of execution costs.” *Journal of Financial Markets* 1, 1-50

Bessembinder, H., and K. Venkataraman. 2010. “Bid-ask Spreads: Measuring Trade Execution Costs in Financial Markets,” in *Encyclopedia of Quantitative Finance*, edited by Rama Cont, John Wiley & Sons.

Biais, Bruno, Thierry Foucault, and Sophie Moinas (2013), “Equilibrium high-frequency trading,” SSRN working paper.

Boyer, Brian H., and Keith Vorkink. 2014. “Stock options as lotteries.” *The Journal of Finance* 69, No. 4, 1485-1527.

Cao, Jie and Bing Han. 2013. “Cross section of option returns and idiosyncratic stock volatility.” *Journal of Financial Economics* 108 No. 1, 231-249.

Cho, Young-Hye, and Robert F. Engle. 1999. “Modeling the Impacts of Market Activity on Bid-Ask Spreads in the Option Market.” *NBER Working Paper Series* No. 7331.

Copeland, Thomas, and Dan Galai. 1983. Information Effects on the Bid-Ask Spread, *Journal of Finance* 38, No. 5, 1457–1469 (December)1983

De Fontnouvelle, Patrick, Raymond P. H. Fishe, and Jeffrey H. Harris, 2003. “The Behavior of Bid-Ask Spreads and Volume in Options Markets during the Competition for Listings in 1999," *Journal of Finance* 58, No. 6, 2437-2464

Doran, James S., Andy Fodor, and Danling Jiang. 2013. "Call-Put Implied Volatility Spreads and Option Returns." *Review of Asset Pricing Studies*.

Driessen, Joost, Pascal J. Maenhout, and Grigory Vilkov. 2009. "The price of correlation risk: Evidence from equity options.” *Journal of Finance* 64, No. 3, 1377-1406.

Engle, Robert, and Breno Neri. 2010. “The impact of hedging costs on the bid and ask spread in the options market.” Working paper, New York University.

George, Thomas J., and Francis A. Longstaff. 1993. “Bid-ask spreads and trading activity in the S&P 100 index options market.” *Journal of Financial and Quantitative Analysis* 28, No. 3.

Goyal, Amit, and Alessio Saretto. "Cross-section of option returns and volatility." Journal of Financial Economics 94.2 (2009): 310-326.

Goyenko, Ruslan, Chayawat Ornthanalai, and Shengzhe Tang. 2014. “Trading Cost Dynamics of Market Making in Equity Options.” Working Paper, McGill University and University of Toronto.

Hasbrouck, Joel. “Measuring the Information Content of Stock Trades.” *The Journal of Finance* 46, no. 1 (1991): 179–207.

Holden, Craig, and Stacey Jacobsen. 2013. “Liquidity measurement problems in fast, competitive markets: expensive and cheap solutions." *Journal of Finance*

Jameson, Mel, and William Wilhelm. "Market making in the options markets and the costs of discrete hedge rebalancing." The Journal of Finance 47.2 (1992): 765-779.

Kaul, Gautam, Mahendrarajah Nimalendran, and Donghang Zhang. 2004. "Informed trading and option spreads." Available at SSRN 547462.

Muravyev, Dmitriy. 2014. “Order Flow and Option Expected Returns. Working paper, Boston College

Muravyev, Dmitriy, Neil D. Pearson and John P. Broussard. 2013. Is there Price Discovery in Equity Options? *Journal of Financial Economics*.

Petrella, Giovanni. 2006. “Option bid-ask spread and scalping risk: Evidence from a covered warrants market.” *Journal of Futures Markets* 26.9 (2006): 843-867.

Rosch, Eleanor. 1975. “Cognitive reference points.” *Cognitive Psychology* 7.4: 532-547.

Figure 1 Stylized example of execution timing

The figure shows how option prices (vertical axis) evolve in time (horizontal axis). The current quote midpoint (grey) eventually converges to the future expected midpoint (green) implied by the current price of the underlying. Traders who time executions wait until the expected quote midpoint approaches the bid price (dark blue) to execute their sell trades (solid blue arrows). Here traders timed their sales well because if they waited a little bit more, the bid price would have decreased, increasing costs for sell trades. Conventional measures of the bid-ask spread and price impact that use the current quote midpoint overestimate trading costs because the quote midpoint is on average significantly higher than its expected value at the time of sell trades, and price will decrease even if no sell trades arrive.

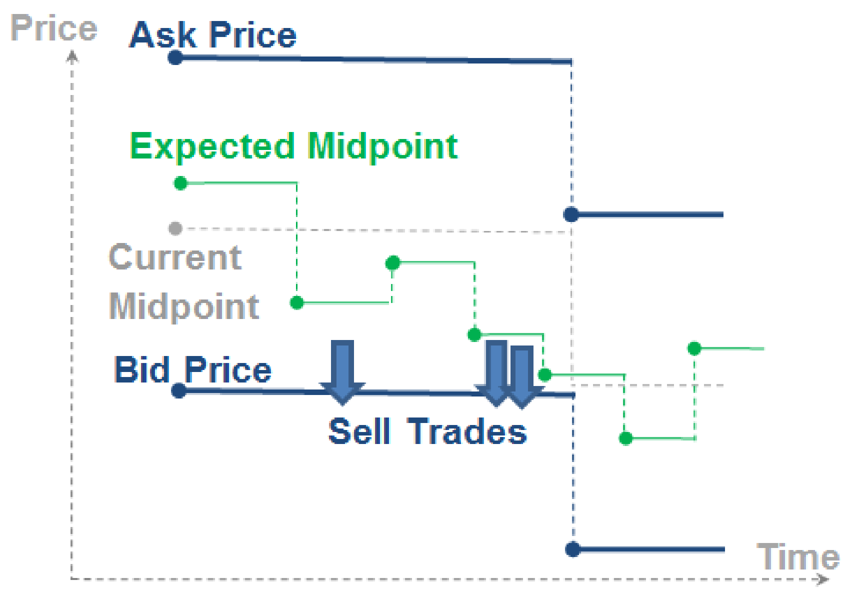

Figure 2 Price changes and expected price changes conditional on signed trade size

The red line shows the predicted change in option price based on public information immediately before a trade computed using Eq. (3) with a 10-minute horizon, as a function of the signed trade size, in dollars. The blue line shows the option price changes during the 10 minutes following a set of simulated trades for the same option and date at random times that do not overlap with the 10 minutes periods following the time of an actual trade. The difference between the red and blue lines is due to execution timing. The green line shows the change in the option price midpoint from the time of a trade until 10 minutes after the trade. It differs from the red line because not all traders time executions and because option prices change due to inventory and adverse-selection impacts.

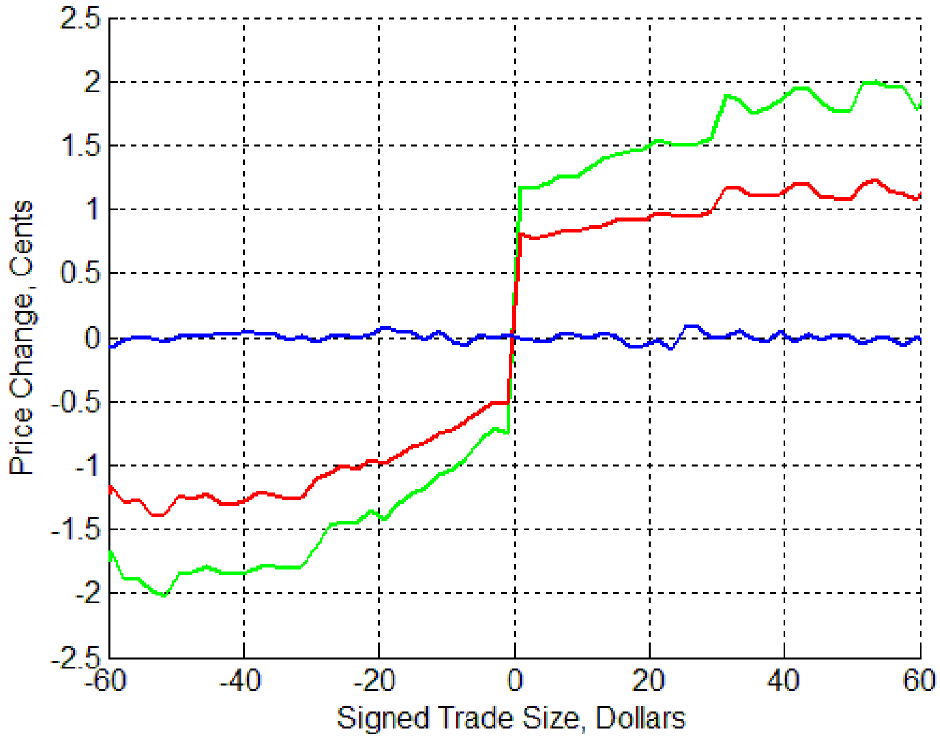

Figure 3 Adoption of algorithmic trading in the options market
The fraction of trades that display execution timing computed using Eq. (5) and the regression model for expected price changes in Eq. (3). Each point on the graph corresponds to a 20-trading-day period. Because it is likely that algorithmic traders sometimes execute trades without timing the executions, the estimates provide lower bounds on the fraction of liquidity-taking trades executed by algorithmic traders.

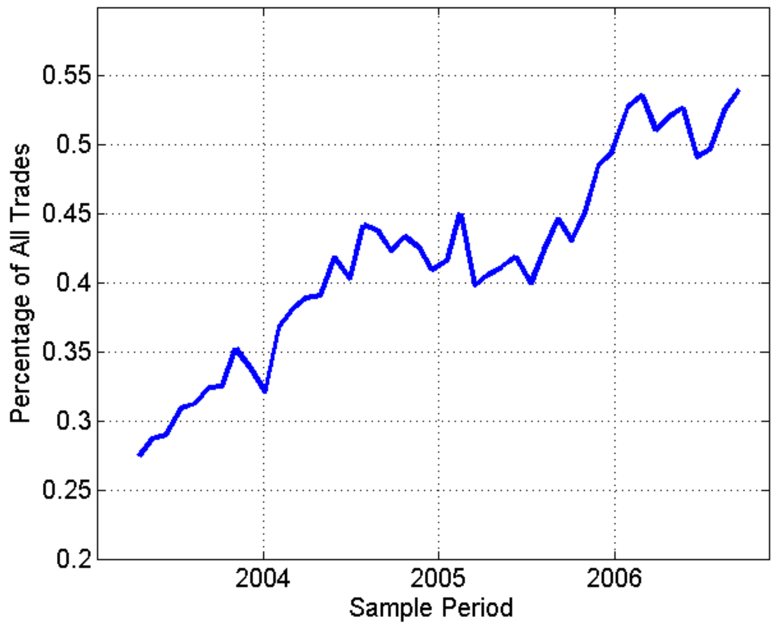

Figure 4 Changes in average quotes, conventional, and adjusted effective spreads over time

The graph shows the evolution of the average quoted (black), the conventionaly measured effective (blue), and the adjusted effective (red) bid-ask spreads over the sample period. The adjusted-effective bid-ask spread is computed as twice the difference between the trade price and the expected future midpoint predicted by regression in Eq. (3). The effective spread is twice the difference between trade price and the quote midpoint. The average quoted spread is computed as an average of one-second quoted spread snapshots over the entire day for an option contract involved in a given trade transaction. Each point on the graph is an average across all option trades on a given day. The sample period is from April 2003 to October 2006.

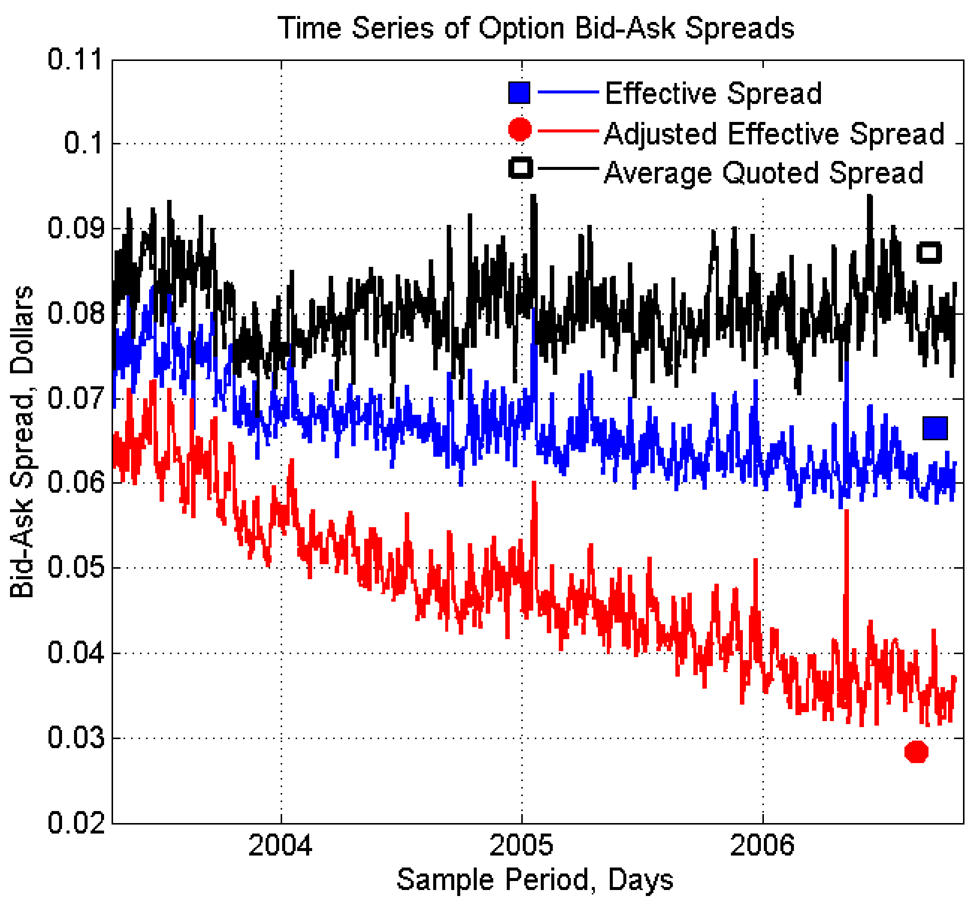

Figure 5 Relations between various measures of the spread and option moneyness as measured by |Δ|

The graph plots non-parametric estimates for five bid-ask spread measures as functions of the absolute of the option delta, |Δ|. The spread measures are the average quoted spread on the day of an option trade (dash-dot black), the quoted spread at the time of the trade (blue), the effective spread (dashed blue), the adjusted effective spread (red line) based on the regression model, and finally, the adjusted effective spread for trades of at least ten contracts (dashed magenta). The adjusted-effective bid-ask spread is computed as twice the difference between the trade price and the expected future midpoint predicted by the regression model in Eq. (3). The effective spread is twice the difference between trade price and the quote midpoint. The average quoted spread is computed as an average of one-second quoted spread snapshots over the entire day for the option contract involved in the trade transaction. The lines are estimated with a kernel regression based on the sample of 20 million trades for options on 39 stocks from April 2003 to October 2006. Option deltas are computed from the Black-Scholes-Merton formula.

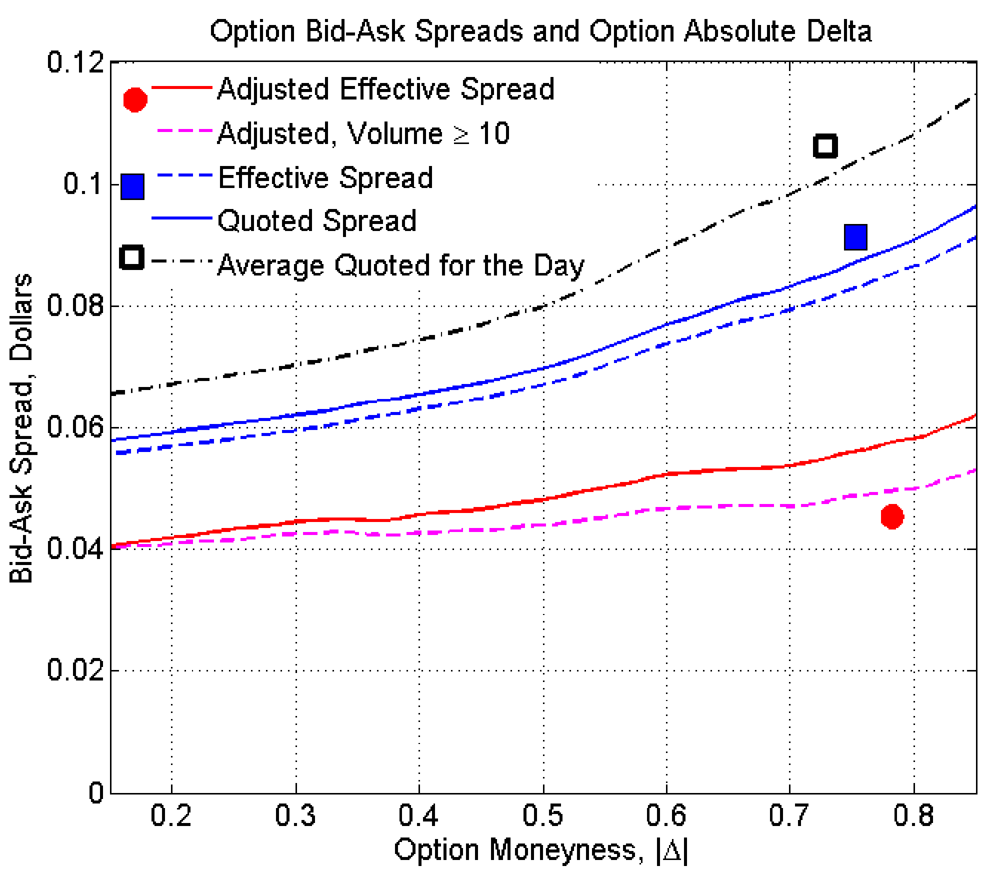

Figure 6 Observed price impacts compared to regression and BSM estimates of expected price changes based on public information

Observed price impact (blue) is measured as the difference between the quote midpoint ten minutes after a trade and the midpoint immediately before it. The expected price change based on the Black-Scholes-Merton model (black) is computed as the difference between the BSM implied option price and the pre-trade quote midpoint. The expected price change implied by regression in Eq. (3) (red) is computed as the difference between the expected future midpoint predicted by Eq. (3) and the pre-trade quote midpoint. The distribution of trade size is highly skewed (roughly exponential). Mean trade size is 42 contracts, and its 50^{th} and 95^{th} percentiles are 10 and 114 contracts respectively. Trade size is reported in contracts, each on 100 underlying shares.

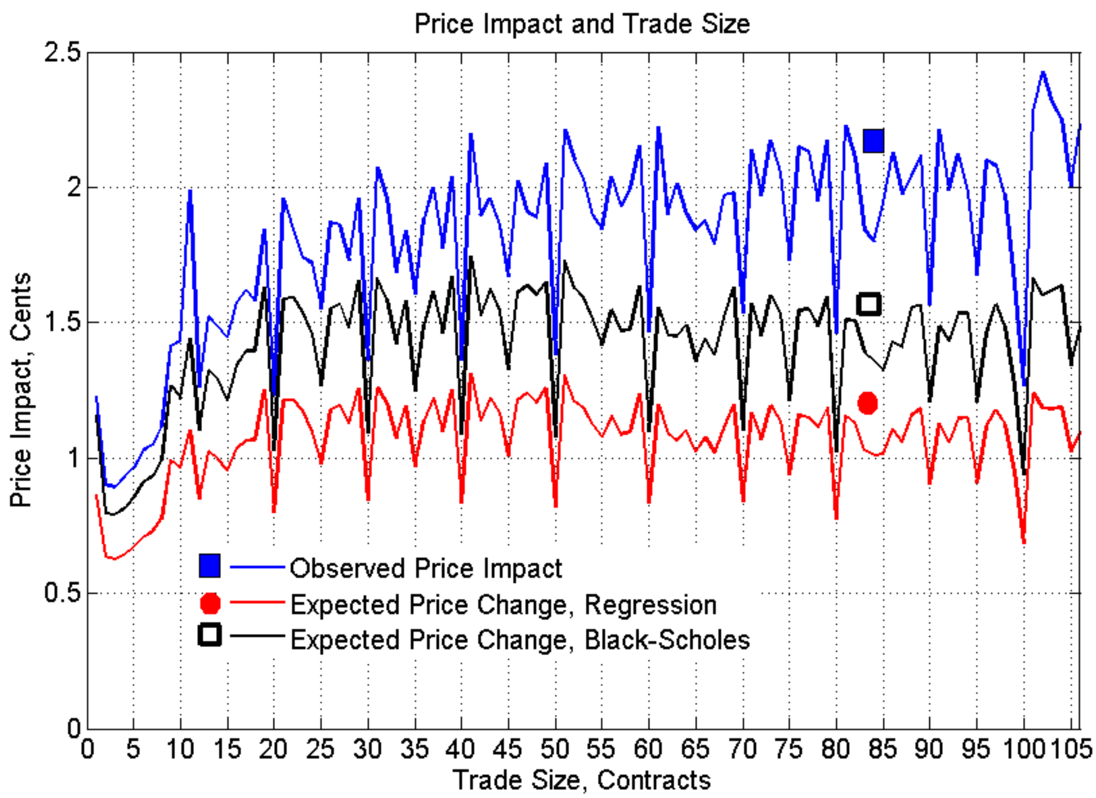

Figure 7 Price impact adjusted for expected price changes based on public information at the time of the trade

Observed price impact is adjusted by the expected changes in the quote midpoint to extract the causal impact of trades. The Black-Scholes-Merton method (blue) assumes that the quote midpoint is expected to converge to the option price implied from the current price for the underlying. The regression method (red) estimates expected changes in price using the regression model in Eq. (3). Each data point on the graph is computed as a simple average across all option trades of a given trade size. The distribution of trade size is highly skewed (roughly exponential). The mean trade size is 10, and its 50th and 95th percentiles are 42 and 114 contracts respectively. Trade size is in contracts each on 100 underlying shares. The smoothing of the BSM-adjusted price impacts (in black) is done using a kernel regression.

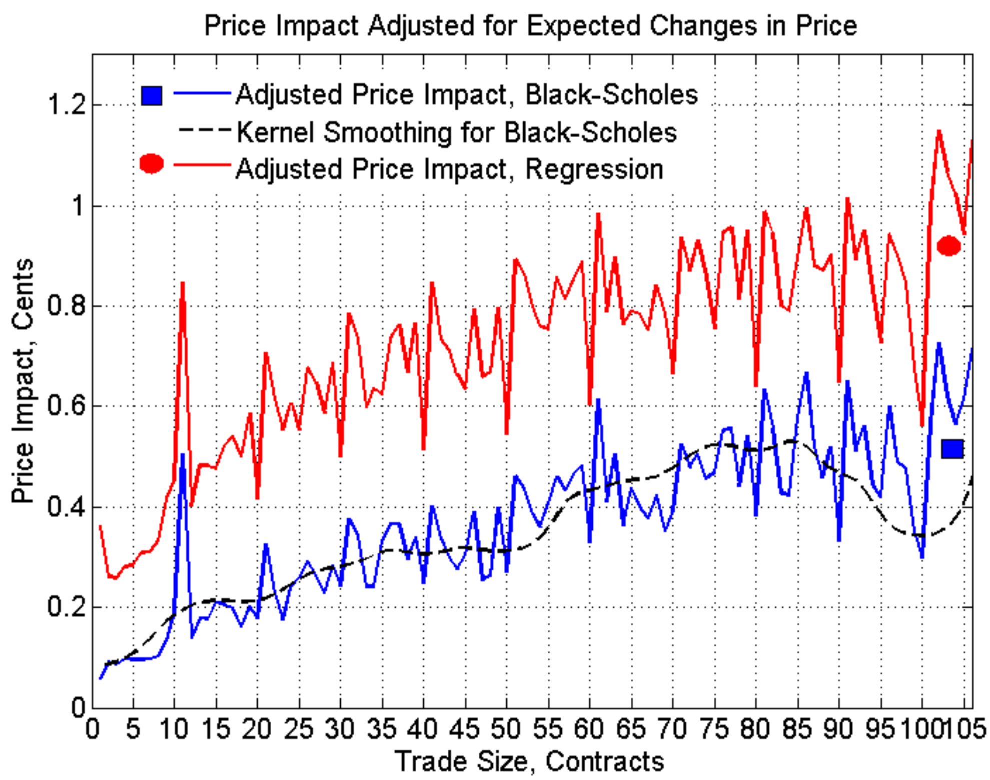

Table 1 Summary statistics

The reported variables include the execution timing bias for the BSM and the regression methods as defined in Eq. (4), the observed price impact and the expected changes in the quote midpoint for the BSM and regression methods, the absolute value of the option delta, the square root of the option time-to-expiration measured in calendar days, a dummy taking the value one for a call option, a time trend, the option price and the bid-ask spread, a dummy variable for buyer initiated trades, and and several measures derived from the option trade size in contracts. Round trades are trades with size divisible by ten and larger than 15. The time trend is in calendar years and normalized to zero at the beginning of the sample period. #ExchAtNBBO is the number of exchanges quoting best price in the direction of a trade. Dummy variables taking the value one if condition *x* is satisfied are denoted I_x. The table reports the mean and standard deviation and 25th, 50th, and 75th percentiles of the distributions of the variables.

Variable	Mean	Std. Dev.	25%	50%	75%
Timing Bias (BSM)	38%	78%	–9%	29%	82%
Timing Bias (regression)	32%	69%	–6%	26%	70%
$\Delta P_t^{t+\tau}$, 10minutes (cents)	1.31	5.80	0.00	0.00	5.00
$\Delta \hat{P}_t^{BSM}$ (cents)	1.10	2.39	-0.29	0.91	2.32
$\Delta \hat{P}_t^{t+\tau}$, 10minutes (cents)	0.84	1.79	-0.15	0.65	1.67
$\lvert\Delta\rvert$	0.45	0.20	0.30	0.44	0.59
$\sqrt{T-t}$	8.57	5.10	5.00	6.71	11.09
D(Call)	0.64	0.48	0.00	1.00	1.00
Time Trend	1.72	1.07	0.77	1.69	2.69
Option Price ($)	1.70	1.94	0.60	1.10	2.10
Bid-Ask (cents)	6.42	4.13	5.00	5.00	10.00
I_{Buy}	0.46	0.50	0.00	0.00	1.00
$\sqrt{\text{Size}}$	4.14	5.03	1.73	3.16	4.47
$I_{\text{Size = 1}}$	0.14	0.35	0.00	0.00	0.00
$I_{\text{Size > 15}}$	0.31	0.46	0.00	0.00	1.00
I_{Round}	0.18	0.39	0.00	0.00	1.00
#ExchAtNBBO	2.97	1.85	1.00	3.00	5.00
$I_{\text{\#Exch = 1}}$	0.34	0.47	0.00	0.00	1.00

Table 2 Prediction of future option price movements using the implied option price
The table reports estimates of the slope coefficient α_1 and the R^2 from regressions

$$P_{t+\tau} - P_t = \alpha_0 + \alpha_1(\hat{P}_t^{BSM} - P_t) + \varepsilon_t$$

of the change in the option quote midpoint on the Black-Scholes-Merton implied bias $\hat{P}_t^{BSM} - P_t$. The regressions are estimated separately for each stock and day using option quote midpoints observed at two minute frequency, and the table reports the time-series averages of the daily coefficient estimates for each stock. Estimates are reported for three time horizons: one minute, ten minutes, and one hour. All coefficient estimates are highly statistically significant, with *t*-statistics for the average coefficient estimates ranging from 5 to 100, with a median value of 29. The intercept is included in the regression model but is not reported.

	1 minute		10 minutes		1 hour	
Ticker	Coeff.	R^2	Coeff.	R^2	Coeff.	R^2
Average	0.54	22%	0.66	10%	0.66	3%
Std. Dev.	0.11	5%	0.10	4%	0.12	2%
AIG	0.46	18%	0.49	4%	0.49	1%
AMAT	0.64	29%	0.77	14%	0.82	5%
AMGN	0.61	23%	0.66	6%	0.58	1%
AMR	0.44	18%	0.53	8%	0.49	2%
AMZN	0.60	22%	0.67	6%	0.64	1%
AOL	0.39	16%	0.62	14%	0.73	6%
BMY	0.37	13%	0.51	7%	0.64	4%
BRCM	0.66	24%	0.69	6%	0.64	1%
C	0.55	26%	0.71	13%	0.71	4%
COF	0.48	17%	0.53	4%	0.37	1%
CPN	0.34	12%	0.50	13%	0.57	7%
CSCO	0.55	24%	0.73	15%	0.76	5%
DELL	0.61	28%	0.70	11%	0.73	3%
EBAY	0.66	23%	0.69	5%	0.63	1%
EMC	0.47	20%	0.65	16%	0.72	7%
F	0.40	15%	0.59	14%	0.67	6%
GE	0.52	24%	0.72	15%	0.75	5%
GM	0.43	18%	0.55	7%	0.52	2%
HD	0.64	30%	0.81	14%	0.78	4%
IBM	0.62	26%	0.73	8%	0.69	2%
INTC	0.54	23%	0.70	13%	0.74	4%
JPM	0.51	23%	0.66	12%	0.73	4%
KLAC	0.69	26%	0.76	6%	0.78	2%
MMM	0.60	24%	0.69	7%	0.63	2%
MO	0.36	14%	0.40	4%	0.34	1%
MSFT	0.50	22%	0.65	15%	0.68	6%
MWD	0.50	19%	0.58	6%	0.47	1%

NXTL	0.55	24%	0.63	9%	0.55	2%
ORCL	0.54	23%	0.69	17%	0.72	7%
PFE	0.41	17%	0.51	7%	0.57	3%
QCOM	0.68	27%	0.74	7%	0.69	2%
QLGC	0.66	26%	0.73	8%	0.70	2%
QQQ	0.47	17%	0.73	11%	0.76	3%
QQQQ	0.69	26%	0.82	13%	0.81	4%
SBC	0.45	19%	0.63	12%	0.66	4%
SMH	0.68	30%	0.81	11%	0.83	3%
TYC	0.45	18%	0.58	10%	0.53	3%
XLNX	0.69	29%	0.77	9%	0.78	2%
XOM	0.62	29%	0.73	9%	0.65	2%

Table 3 Expected changes in the option quote midpoint

Regression of changes in option quote midpoint for ten minutes and one hour horizons on the explanatory variables as well as lagged changes in the option and delta-adjusted stock quote midpoints.

$$P_{t+\tau} - P_t = \alpha_0 + \alpha_1(\hat{P}_t^{BSM} - P_t) + \alpha_2(\hat{P}_t^{BBO} - P_t) + \alpha_3\#\text{ExchBid} + \alpha_4\#\text{ExchAsk} + \sum_{i=1}^{12}\alpha_{i+4}(\Delta_{t-i\Delta t} dS_{t-i\Delta t}) + \sum_{i=1}^{12}\alpha_{i+16} dP_{t-i\Delta t} + \varepsilon_t$$

The explanatory variables include the BSM implied bias (the difference between the BSM-implied price and the NBBO quote midpoint), the difference between the average quote midpoint across all exchanges and the current NBBO quote midpoint (Average BBO Price – NBBO), and number of exchanges at the best ask and best bid. The regression is estimated separately for each stock and six absolute delta (0.35 and 0.65 cut-offs) and time-to-expiration (60 days cut-off) bins within each day, average coefficients are reported. The lagged quote changes are based on twelve regularly spaced five-second time periods (only the first two and the sum of all twelve coefficients are reported). The regressions are based on the pooled sample of option quote midpoint snapshots with two minute time step. All price changes are measured in cents.

Days-to-Expiration	Money-ness	Inter-cept	BSM Implied Bias	Average BBO – NBBO	# Exchs. at Bid	# Exchs. at Ask	Stock price changes adjusted for option delta			Changes in option quote midpoint			
							$t-1$	$t-2$	Sum t–1 through t–12	$t-1$	$t-2$	Sum t–1 through t–12	R^2
τ = 10 minutes													
short-term	OTM	-0.02	0.26	0.27	0.02	-0.01	0.38	0.24	1.84	-0.17	-0.15	-1.22	0.17
long-term	OTM	-0.03	0.33	0.13	0.07	-0.07	0.39	0.25	1.90	-0.20	-0.18	-1.49	0.13
short-term	ATM	-0.08	0.40	0.31	0.05	-0.05	0.39	0.24	1.78	-0.14	-0.13	-1.15	0.12
long-term	ATM	-0.09	0.44	0.17	0.12	-0.11	0.40	0.25	1.91	-0.16	-0.15	-1.30	0.12
short-term	ITM	-0.27	0.54	0.23	0.09	-0.07	0.32	0.19	1.28	-0.12	-0.12	-1.06	0.09
long-term	ITM	-0.17	0.51	0.18	0.13	-0.11	0.32	0.19	1.34	-0.14	-0.14	-1.22	0.11
τ = 1 hour													
short-term	OTM	-0.04	0.31	0.39	0.11	-0.09	0.27	0.13	0.66	-0.07	-0.07	-0.63	0.18
long-term	OTM	-0.08	0.42	0.29	0.12	-0.11	0.25	0.11	0.46	-0.09	-0.09	-0.77	0.13
short-term	ATM	-0.15	0.45	0.44	0.07	-0.06	0.24	0.10	0.33	-0.08	-0.08	-0.82	0.10
long-term	ATM	-0.16	0.53	0.32	0.10	-0.08	0.23	0.09	0.26	-0.08	-0.08	-0.74	0.08
short-term	ITM	-1.03	0.63	0.33	0.06	-0.02	0.13	0.01	-0.42	-0.05	-0.06	-0.69	0.08
long-term	ITM	-0.53	0.56	0.35	0.02	0.02	0.15	0.04	-0.27	-0.10	-0.10	-1.02	0.08

Table 4 Evolution of trading costs of investors who do and do not use execution timing

The costs are measured by the adjusted effective bid-ask spread in Panel A and by the conventional effective spread in Panel B. The adjusted effective spread is twice the difference between the trade price and the expected future quote midpoint (estimated by the regression model in Eq. 3 with a ten-minute horizon) and then adjusted for trade direction, i.e. $2I_{B/S}(TP_t - \hat{P}_{t+T})$. The effective spread is twice the absolute difference between the trade price and the midpoint immediately prior to the trade, i.e., $2|TP_t - P_t|$. The first three columns report adjusted effective spreads for traders who time executions ("Timers") and those who do not ("Non-Timers") for the first and last 20-day periods (months) during the sample period as well as the average across all months. Each trade is classified as either initiated by an execution timer or by a non-timer. The last column reports the percentage of total trades initiated by algorithms computed using Eq. (5). All variables, except the percentage of algorithmic trades, are reported in cents.

Panel A Adjusted Effective Spreads

	Adjusted Effective Spread, Cents			Percentage of Trades Initiated by Timers
	All	Timers	Non-Timers	
First Month	5.5	1.9	6.8	27.5%
Last Month	3.5	1.1	6.0	54.0%
All Months	4.3	1.3	6.2	40.5%

Panel B Conventional Effective Spreads

	Effective Spread, Cents			Percentage of Trades Initiated by Timers
	All	Timers	Non-Timers	
First Month	6.8	6.7	6.8	27.5%
Last Month	6.0	6.0	6.0	54.0%
All Months	6.2	6.2	6.2	40.5%

Table 5 Option bid-ask spread measures by stock
The average quoted spread is computed as an average of one-second quoted spread snapshots over the entire day for an option contract involved in a given trade transaction. The effective spread is twice the difference between trade price and the quote midpoint. The adjusted-effective bid-ask spread is computed as twice the difference between the trade price and the expected future midpoint predicted by regression in Eq. (3). The black-Scholes adjusted spread is twice the difference between trade price and option price implied from the price of the underlying and the lagged implied volatility as in Eq. (1). The last column reports the ratio between the adjusted effective spread and the average quoted spread. The average quoted spread significantly overestimates the costs of taking liquidity. All spreads are reported in cents. Number of observations for each stock is in thousands. An average and standard deviation across 39 stocks are reported at the top.

Ticker	# of obs., in 1,000s	Bid-Ask Spread, Cents						Ratio of Adjusted to Avg. Quoted Spread
		Avg. Quoted	Quoted at time of trade	Effec -tive	Reali-zed	Adjusted, BSM	Adjust-ed, Reg-ression	
Average		8.4	6.6	6.4	3.7	4.2	4.5	54%
Std.Dev.		1.6	1.0	1.0	0.5	0.7	0.7	
AIG	319	10.6	8.1	7.8	4.2	5.1	5.5	52%
AMAT	414	7.0	5.6	5.3	3.3	3.6	3.9	56%
AMGN	551	10.2	7.8	7.4	4.0	4.4	4.8	47%
AMR	295	9.3	7.0	6.7	3.1	4.2	4.5	48%
AMZN	671	9.3	7.0	6.7	3.4	3.8	4.3	46%
AOL	52	7.0	6.0	6.0	4.7	5.1	5.1	73%
BMY	220	7.5	6.1	5.9	3.6	4.3	4.6	61%
BRCM	552	9.7	7.3	6.9	3.4	3.8	4.3	44%
C	475	8.0	6.5	6.3	4.1	4.5	4.9	61%
COF	182	12.1	9.0	8.5	4.1	5.4	5.8	48%
CPN	115	7.4	6.1	5.9	3.9	4.6	4.8	65%
CSCO	775	6.4	5.3	5.2	3.4	3.5	3.8	59%
DELL	506	7.5	6.0	5.8	3.4	3.6	4.0	53%
EBAY	1,246	10.1	7.8	7.4	4.0	4.3	4.9	49%
EMC	251	6.8	5.4	5.3	3.5	3.8	3.9	57%
F	203	6.9	5.7	5.6	3.6	4.2	4.4	64%
GE	614	6.8	5.7	5.6	3.9	4.2	4.4	65%
GM	619	10.1	7.8	7.5	3.6	4.8	5.0	50%
HD	372	8.1	6.5	6.3	3.9	4.4	4.7	58%
IBM	740	9.7	7.6	7.3	4.3	4.8	5.2	54%
INTC	1,228	6.4	5.4	5.2	3.4	3.5	3.8	59%
JPM	380	8.4	6.8	6.5	4.2	4.7	5.0	60%
KLAC	313	9.8	7.3	6.9	3.5	3.8	4.1	42%
MMM	274	11.5	8.8	8.4	4.7	5.6	6.0	52%

MO	572	10.4	8.2	7.9	4.4	5.4	5.8	56%
MSFT	970	6.4	5.4	5.2	3.4	3.5	3.8	59%
MWD	148	9.9	7.6	7.3	4.0	5.0	5.3	54%
NXTL	192	8.2	6.5	6.4	4.2	4.7	5.0	61%
ORCL	306	6.4	5.3	5.1	3.3	3.5	3.7	58%
PFE	711	7.0	5.8	5.7	3.7	4.0	4.3	61%
QCOM	750	9.2	7.0	6.8	3.6	3.9	4.2	46%
QLGC	243	9.8	7.3	6.9	3.5	4.0	4.3	44%
QQQ	1,963	6.3	5.4	5.2	3.7	3.7	3.9	62%
QQQQ	1,840	6.2	5.0	4.7	2.2	2.3	2.6	42%
SBC	108	7.6	6.1	6.0	4.3	4.7	4.9	64%
SMH	387	7.9	6.0	5.7	2.9	3.5	3.7	47%
TYC	225	8.4	6.7	6.4	4.1	4.7	4.9	58%
XLNX	170	8.8	6.4	6.1	3.2	3.7	3.9	44%
XOM	537	8.7	6.9	6.7	3.6	4.2	4.7	54%

Table 6 Price impacts and expected changes in the quote midpoint
Observed price impact is measured as the difference between the quote midpoint immediately before a trade and ten (one) minutes later, $P_{t+T} - P_t$. The expected price change is the difference between the expected future midpoint predicted by regression in Eq. (3) and the pre-trade quote midpoint, $\hat{P}_{t+T} - P_t$. The adjusted price impact is the difference between observed impact and the expected change in price, $\Delta P_{t+T} - \Delta\hat{P}_{t+T}$. It measures the causal impact a trade. The last column reports the ratio between the observed and adjusted price impacts. Observed price impact significantly overestimates the causal impact of trades. The price changes are reported for one and ten minute horizons and are measured in cents. An average and standard deviation across 39 stocks are reported at the top.

Ticker	Observed Price Impact, Cents		Expected Price Change, Cents		Adjusted Price Impact, Cents		Ratio of Adjusted to Observed Impact
	1 Minute	10 Minutes	1 Minute	10 Minutes	1 Minute	10 Minutes	
Average	1.13	1.34	0.47	0.82	0.66	0.52	39%
Std.Dev.	0.37	0.40	0.24	0.30	0.19	0.18	
AIG	1.51	1.80	0.64	1.08	0.87	0.72	40%
AMAT	0.87	1.01	0.36	0.66	0.51	0.35	35%
AMGN	1.52	1.73	0.81	1.21	0.71	0.52	30%
AMR	1.45	1.79	0.42	0.87	1.03	0.92	51%
AMZN	1.42	1.64	0.76	1.16	0.66	0.48	29%
AOL	0.48	0.65	0.13	0.31	0.35	0.34	52%
BMY	0.92	1.12	0.24	0.53	0.68	0.59	53%
BRCM	1.64	1.77	0.90	1.31	0.74	0.46	26%
C	0.96	1.13	0.34	0.66	0.62	0.47	42%
COF	1.82	2.26	0.76	1.24	1.06	1.02	45%
CPN	0.81	1.00	0.11	0.32	0.70	0.68	68%
CSCO	0.74	0.93	0.26	0.57	0.48	0.36	39%
DELL	1.06	1.23	0.45	0.81	0.61	0.42	34%
EBAY	1.59	1.76	0.88	1.25	0.71	0.51	29%
EMC	0.74	0.93	0.21	0.51	0.53	0.42	45%
F	0.78	0.98	0.15	0.42	0.63	0.56	57%
GE	0.67	0.84	0.22	0.46	0.45	0.38	45%
GM	1.61	1.95	0.47	1.00	1.14	0.95	49%
HD	1.03	1.19	0.39	0.72	0.64	0.47	39%
IBM	1.34	1.53	0.63	1.00	0.71	0.53	35%
INTC	0.75	0.95	0.31	0.62	0.44	0.33	35%
JPM	0.94	1.18	0.33	0.66	0.61	0.52	44%
KLAC	1.61	1.75	0.97	1.31	0.64	0.44	25%
MMM	1.64	1.87	0.72	1.13	0.92	0.74	40%
MO	1.48	1.78	0.52	0.93	0.96	0.85	48%

MSFT	0.77	0.95	0.25	0.57	0.52	0.38	40%
MWD	1.41	1.64	0.55	0.92	0.86	0.72	44%
NXTL	0.96	1.11	0.36	0.66	0.60	0.45	41%
ORCL	0.75	0.94	0.23	0.53	0.52	0.41	44%
PFE	0.83	1.02	0.26	0.56	0.57	0.46	45%
QCOM	1.48	1.59	0.82	1.21	0.66	0.38	24%
QLGC	1.52	1.72	0.74	1.18	0.78	0.54	31%
QQQ	0.56	0.79	0.29	0.58	0.27	0.21	27%
QQQQ	1.01	1.23	0.48	0.91	0.53	0.32	26%
SBC	0.68	0.85	0.20	0.43	0.48	0.42	49%
SMH	1.21	1.41	0.54	0.94	0.67	0.47	33%
TYC	0.99	1.20	0.32	0.64	0.67	0.56	47%
XLNX	1.35	1.50	0.66	1.04	0.69	0.46	31%
XOM	1.33	1.54	0.51	0.92	0.82	0.62	40%

Table 7 Price impact regressions explaining changes in the quote midpoint $\Delta P_t^{t+\tau}$ and the BSM-adjusted change $\Delta P_t^{t+\tau} - \Delta \hat{P}_t^{BSM}$

In the first three specifications the measure of price impact is the change in the quote midpoint $\Delta P_t^{t+\tau} = P_{t+\tau} - P_t$ during the ten minutes after a trade at time *t*. The fourth regression uses the BSM-adjusted price impact $\Delta P_t^{t+\tau} - \Delta \hat{P}_t^{BSM}$ as the dependent variable. The right-hand side variables include the absolute value of the option delta (|Δ|), the square root of the time to option expiration ($\sqrt{T-t}$), a dummy variable taking the value of one if the option traded is a call (I_{Call}), a time trend measured in years (TimeTrend), the option bid-ask spread measured in cents (Bid-Ask), a dummy variable taking the value of one if the trade is a purchase (I_{Buy}), the square root of the trade size measured in contracts ($\sqrt{\text{Size}}$), a dummy variable taking the value one if the trade size is one contract ($I_{\text{Size = 1}}$), the number of option exchanges at the NBBO on the side of the market where the trade occurred (#ExchAtNBBO), a dummy variable taking the value one if there is only one exchange at the NBBO on the side of the market where the trade occurred ($I_{\text{\#Exch = 1}}$). The second and third specifications also include the predicted price change from the regression model, $\Delta \hat{P}_t^{t+\tau}$, and the fourth specification allows for a non-linear relation between the predicted and actual prices changes by including dummy variables $\Delta \hat{P}_t^{t+\tau} I_{0<x\leq 2}$, $\Delta \hat{P}_t^{t+\tau} I_{2<x\leq 5}$, and $\Delta \hat{P}_t^{t+T} I_{5<x}$ that take the value one if the predicted quote change l is between zero and two cents, two and five cents, and greater than five cents, respectively. *t*-statistics based on robust standard errors, which are clustered by date, are reported in parentheses. Stock fixed effects are included in the regressions but not reported.

	$\Delta P_t^{t+\tau}$	$\Delta P_t^{t+\tau}$	$\Delta P_t^{t+\tau}$	$\Delta P_t^{t+\tau} - \Delta \hat{P}_t^{BSM}$
$\Delta \hat{P}_t^{t+\tau}$ (cents)		0.776	0.584	
		(66.92)	(32.60)	
|Δ|	0.102	-0.127	-0.315	0.201
	(4.20)	(5.77)	(13.53)	(8.87)
$\sqrt{T-t}$	-0.030	-0.005	-0.005	0.008
	(33.51)	(5.68)	(5.35)	(10.93)
I_{Call}	-0.018	-0.046	-0.041	-0.079
	(2.86)	(7.25)	(6.50)	(11.28)
TimeTrend	0.552	0.228	0.201	0.043
	(92.37)	(33.31)	(34.23)	(9.75)
Option Price ($)	0.158	0.079	0.082	0.010
	(38.50)	(20.01)	(19.82)	(2.65)
Bid-Ask (cents)	0.061	0.017	0.012	0.008
	(26.70)	(7.74)	(6.08)	(4.97)
I_{Buy}	0.116	0.144	0.145	0.087
	(7.88)	(8.98)	(9.08)	(4.43)

$\sqrt{\text{Size}}$	0.019	0.019	0.018	0.018
	(50.26)	(50.22)	(50.43)	(48.73)
$I_{\text{Size=1}}$	-0.160	-0.079	-0.076	-0.058
	(16.43)	(8.51)	(8.17)	(6.06)
#ExchAtNBBO	-0.571	-0.198	-0.162	-0.015
	(139.13)	(27.08)	(28.51)	(5.11)
$I_{\text{\#Exch = 1}}$	0.562	0.194	0.155	-0.006
	(35.07)	(19.41)	(16.16)	(0.65)
$\Delta\hat{P}_t^{t+\tau} I_{0<x\leq 2}$			0.435	
			(22.46)	
$\Delta\hat{P}_t^{t+\tau} I_{2<x\leq 5}$			0.366	
			(21.90)	
$\Delta\hat{P}_t^{t+\tau} I_{5<x}$			0.201	
			(15.49)	
R^2	0.05	0.09	0.09	0.00
N (1,000s)	20,483	20,483	20,483	20,483

Table 8 Regressions explaining the execution timing bias

Each column reports the results of a regression of the timing bias on the absolute value of the option delta (|Δ|), the square root of the time to option expiration ($\sqrt{T-t}$), a dummy variable taking the value of one if the option traded is a call (I_{Call}), a time trend measured in years (TimeTrend), the option bid-ask spread measured in cents (Bid-Ask), a dummy variable taking the value of one if the trade is a purchase (I_{Buy}), the square root of the trade size measured in contracts ($\sqrt{\text{Size}}$), a dummy variable taking the value one if the trade size is one contract ($I_{\text{Size = 1}}$), the number of option exchanges at the NBBO on the side of the market where the trade occurred (#ExchAtNBBO), a dummy variable taking the value one if there is only one exchange at the NBBO on the side of the market where the trade occurred ($I_{\text{\#Exch = 1}}$). The last column uses the subsample of trades with more than one exchange quoting the relevant best price. The timing bias for an option is defined in Eq. (4) as the expected one hour change in the quote midpoint from the regression model in Eq. (3) divided by the average quoted spread for the option. *t*-statistics based on robust standard errors, which are clustered by date, are reported in parentheses. Stock fixed effects are included in the regressions but not reported.

Execution Timing (%)	Full Sample	Full Sample	Full Sample
|Δ|	19.367	19.492	0.412
	(46.11)	(45.88)	(0.90)
$\sqrt{T-t}$	–0.009	–0.011	–1.133
	(0.58)	(0.74)	(63.15)
I_{Call}	-2.028	-2.189	1.081
	(16.01)	(18.17)	(11.44)
TimeTrend	6.720	6.743	16.450
	(54.83)	(55.48)	(102.42)
Option Price ($)	4.068	4.097	0.512
	(77.74)	(77.84)	(7.38)
Bid-Ask (cents)	-3.553	-3.570	1.578
	(97.72)	(96.72)	(35.88)
I_{Buy}		-3.564	-6.710
		(9.05)	(20.03)
$\sqrt{\text{Size}}$		-0.160	0.047
		(23.30)	(10.49)
$I_{\text{Size=1}}$		1.843	-2.696
		(6.66)	(14.90)
#ExchAtNBBO			-19.552
			(284.99)
$I_{\text{\#Exch = 1}}$			15.500
			(41.67)
R^2	0.05	0.06	0.33
N (1,000s)	20484	20483	20483

Table 9 Comparison of the conventional and adjusted price impacts for trades of round and non-round sizes

The table reports the results of three regressions explaining different measures of price impact. The first column uses the conventional observed price impact, measured as the dollar difference between the quote midpoint ten minutes after a trade and the pre-trade quote midpoint, $P_{t+\tau} - P_t$. The second column uses the regression-adjusted price impact, computed as the difference between the quote midpoint ten minutes after a trade and the expected future midpoint predicted by the regression model in Eq. (3), $\hat{P}_{t+\tau} - P_t$. The last column uses the BSM-adjusted price impact, computed as the difference between the quote midpoint ten minutes after a trade and the option price implied from the price of the underlying (Eq. 1) immediately before a trade. Each observation is an average of price impacts for a given trade size. Round trades are those for which the trade size is greater than 15 and divisible by 10. Round-five trades are those for which the trade size is greater than 15 and divisible by 5, and are not also round trades. A dummy variable for trade size greater than fifteen contracts $I_{\text{Size}>15}$ is included so that the round-size dummies estimate proper conditional means. Only trade sizes of less than hundred contracts are included. Robust *t*-statistics are reported in parentheses.

	Price Impact (cents)		
	Conventional Price Impact $\Delta P_t^{t+\tau}$	Regression-Adjusted Impact $\Delta P_t^{t+\tau} - \Delta \hat{P}_t^{t+\tau}$	BSM-Adjusted Impact $\Delta P_t^{t+\tau} - \Delta \hat{P}_t^{BSM}$
$I_{\text{Round-five}}$	-0.199	-0.045	0.031
	(6.09)	(2.24)	(1.68)
I_{Round}	-0.414	-0.126	-0.006
	(12.45)	(6.25)	(0.31)
$I_{\text{Size}>15}$	0.156	0.016	-0.041
	(1.78)	(0.34)	(1.21)
Size	-0.017	-0.006	-0.002
	(2.62)	(1.69)	(0.73)
$\sqrt{\text{Size}}$	0.328	0.165	0.104
	(3.23)	(3.02)	(2.90)
Option Price ($)	1.029	0.731	0.705
	(2.89)	(3.73)	(5.30)
R^2	0.85	0.83	0.76
N	99	99	99