Internet appendix to
Order Flow and Expected Option Returns

A.1. Introduction

This appendix reports additional results that supplement the results in the main paper. Specifically, Section A.2 introduces a method to account for endogeneity between trades and quotes and to compute expected price changes; Section A.3 studies the effect of price discreetness and tick size on the price impact components; Section A.4 examines the effect of trade size misclassification; Section A.5 checks for "hot-potato" trading in options; Section A.6 reports how price impact components vary by option exchange; Section A.7 examines outliers and their effect; Section A.8 compares information impact of option trades for individual stocks and ETFs; Section A.9 shows that past order imbalance measures primarily future inventory shocks; Section A.10 discusses the most significant predictors of future option returns excluding past order imbalance; finally, Section A.11 checks the robustness of option return predictability by past order imbalances.

A.2. Accounting for expected changes in price

This section explains how my microstructure method accounts for the effect of endogeneity between trades and quotes on the asymmetric-information component of price impact. Stale public information is a good predictor of future intraday changes in the bid and ask prices. This predictability does not affect the inventory risk component in Eq. (10) but it affects the asymmetric information component in Eq. (9). Trades and quotes are endogenous; more specifically, buy (sell) trades tend to arrive when the quoted price is about to increase (decrease) anyway, so that only a part of the observed increase in price is caused by a trade.

Hasbrouck (1991) and numerous subsequent papers emphasize the importance of accounting for expected price changes in estimating price impact. Muravyev and Pearson (2014) adopt this idea and show its importance for the options market. They find that observed price impact significantly overstate the causal impact of trades if not accounted for expected changes in price. My approach closely follows Muravyev and Pearson (2014).

The expected quote changes due to slow public information diffusion $E(\Delta p_{t*}^{a}|F_{t*})$ at trade times t*, can be estimated in two steps. First, a linear regression approximates $E(\Delta p_{t}^{a}|F_{t})$ and is estimated on historical data for regularly spaced time intervals t. After that, the estimated model is applied to public information F_{t*} at the time of trades.

The first step is implemented as follows. The change in the option ask (bid) price for a given exchange over the next five seconds (to match the evaluation period for price impacts) is predicted by a battery of explanatory variables including short-term price history and the quote deviation from the midpoint. Option and delta-adjusted stock price changes are taken for 12 five-second snapshots to accommodate the most recent price dynamics. The quote deviation from the midpoint is represented by the difference between the quoted ask price and the average quote midpoint across all exchanges[1]. It can also be considered as a measure of the bid-ask spread. If the ask price is close to (far from) the midpoint, then the ask price is likely to increase (decrease) converging to its average. These are arguably the most important variables spanning the available public information F_t; however, other variables may potentially improve the predictability. Because the price impact decomposition only uses exchanges that quote NBBO prices at the time of trade, the regression for the expected quote changes applies the same filter.

$$\Delta p_{t,i}^{a} = \alpha_0 + \alpha_1\left(p_{t,i}^{a} - \mu_t{}^{BBO}\right) + \sum_{n=1}^{12} \alpha_{n+1}(\Delta_{t-n\Delta t} \cdot \Delta S_{t-n\Delta t}) + + \sum_{n=1}^{12} \alpha_{n+13} \Delta p_{t-n\Delta t}^{a} + \epsilon_{t,i} \tag{A1}$$

The regression is estimated separately for each stock and six absolute delta (0.40 and 0.60 cut-offs) and time-to-expiration (70 days cut-off) bins for bid and ask.[2] The average coefficients across all days within each bin are then used for prediction.

Table A4 reports average regression coefficients across all stocks for the five-second time horizon.[3] Changes in the option quote prices are highly predictable with R-square of about 4%,

[1] For the bid price, the difference is reversed, i.e., the BBO mean minus the bid price $\mu_t{}^{BBO} - p_{t,i}^{b}$.

[2] As price dynamics on each day is relatively independent, this methodology simplifies the computation of t-statistics and spotting outliers.

[3] The regression coefficients for 10 second and 1 minute horizons are not reported to save space and can be approximated by multiplying the 5-second coefficients by 1.7 and 5.5 respectively.

and the coefficient signs go in the expected direction. The quote deviation is negative as bid and ask prices converge to their average distance from the midpoint. Consistent with Muravyev, et al. (2013), the option market lags slightly behind the underlying stock; and option price is mean-reverting e.g. because of aggressive limit orders. The intercept is positive for the ask price and negative for the bid because if a market maker is already quoting NBBO price, there is little room for improving it. All the estimates are highly statistically significant and do not vary much across moneyness and time to expiration.

In the next step, the same regression covariates are computed immediately before each option trade and are multiplied by corresponding regression coefficients to compute the expected quote changes. Table 2 summarizes the average expected quote changes for each stock after adjusting for trade direction. Quotes are expected to change in the trade direction by 0.08%. Thus failing to account for the expected quote changes would overstate the information price impact at 0.3% instead of 0.22%. Inventory-risk impact will still be larger the information impact in this case. The expected change estimates are positive for every stock in the sample and range from 0.04% for America Online to 0.13% for QQQQ Nasdaq ETF.

Overall, the intraday dynamics of option quotes is highly predictable; the effect of this predictability on price impacts is significant and need to be accounted for.

A.3. Price discreteness

This section shows that price discreetness does not affect the estimates of the inventory-risk and asymmetric-information price impacts but introduces significant skewness in their sample distributions. This skewness is created because quoted prices do not change in response to most trades if the tick size is large. Because of the skewness, the price impact components should be estimated as an average (and not as a median) over large number of trades.

In practice, prices must take value from a discrete grid with a fixed step (tick size). For example, the US equity market has a tick size of a penny, and thus a price of $10.005 cannot be quoted. During the sample period, the tick size is 5 cents for options with price below three dollars and 10 cents above that price. This tick size is large compared to an average option price of 1.5 dollars.

A large literature studies how market-makers set their quotes in a market with discrete prices.[4] Most theories imply that market makers will widen the quoted bid-ask spread by quoting the nearest above price on the grid for the ask price and the nearest lower price for the bid (i.e., nearest to market-maker's internal bid and ask). For example, if an option market-maker has internal bid/ask prices of 1.38/1.41, she will quote 1.35/1.45 because of a five-cent tick size. Following Hasbrouck (1999), this intuition can be summarized in the following equations. If A_t^* and B_t^* are the bid and ask prices market-maker would quote if prices were continuous, and $\{K, 2K, 3K, \ldots\}$ is the grid of allowed prices, where K is tick size, then the observed bid and ask prices projected on the grid are:

$$A_t = Ceiling\left(\frac{A_t^*}{K}\right) * K; \quad B(t) = Floor\left(\frac{B_t^*}{K}\right) * K \tag{A2}$$

I follow this simple approach to introduce price discreetness in the baseline model of Section 3.2.

Price discreetness has a minor effect on most microstructure methods; however, Harris (1990) and Dravid (1991) show that it may affect stock returns and volatility. Therefore, its effect on the price impact decomposition should be studied. I estimate this effect by conducting numeric simulations with parameters set to match the statistics of my option sample. In each iteration, a trade arrives and quote prices at the trading and non-trading exchanges respond to it, the response includes information and inventory price impacts and the error term as in Eq. (7-7'). These internal prices before and after the trade are projected on the discrete grid of observed prices following Eq. (A2). I then estimate the price impact components for each simulated trade as if these discrete prices were actual data. These individual estimates are then averaged over a large number of simulated trades to produce final estimates for the price impact components. These estimates can be then compared to their true values of information and inventory impacts (which are known in the simulation) to confirm that price discreetness indeed does not introduces any bias.

Specifically, the simulation procedure follows section Section 3.1 and considers the simplest case with one non-trading and one trading exchanges quoting the same ask price, no endogeneity between prices and trades and price impacts are computed in dollar rather than relative terms. These features can be easily added but the distributions then will have both discrete and continuous parts, which make them less intuitive. In each iteration, quoted prices

[4] Hasbrouck (1999), Kandel and Marx (1997), Chordia and Subrahmanyam (1995), Glosten (1994) among others

respond to a single buy trade. Two market-makers have the same before-trade internal ask price $p^*_{0,tr} = p^*_{0,n-tr}$, which is chosen at random to match the distribution of option prices in the data. The observed price $p_{0,tr}$ is discrete and is computed according to equation (A2) with tick size of five cents (options tick size): $p_{0,tr} = Discr_{0.05}(p^*_{0,tr})$. Following Eq. (7-7'), the after-trade ask prices are computed as

$$p^*_{1,tr} = p^*_{0,tr} + InfPI + InvPI + \epsilon_{Inf} + \epsilon_{Inv} \quad \text{(A3)}$$

for the trading exchange, and

$$p^*_{1,n-tr} = p^*_{0,n-tr} + InfPI + \epsilon_{Inf} \quad \text{(A3')}$$

for the non-trading exchange, where the price impacts and error terms are set to match summary statistics of the data (Table 1), the error term ϵ is assumed to be normally distributed.[5] These internal price responses are projected on the discrete grid of observed prices:

$$\Delta p_{tr} = p_{1,tr} - p_{0,tr} = Discr_{0.05}\left(p^*_0 + InfPI + InvPI + +\epsilon_{Inf} + \epsilon_{Inv}\right) - Discr_{0.05}(p^*_0) \quad \text{(A4)}$$

$$\Delta p_{n-tr} = p_{1,n-tr} - p_{0,n-tr} = Discr_{0.05}\left(p^*_0 + InfPI + \epsilon_{Inf}\right) - Discr_{0.05}(p^*_0)$$

Then the information and inventory impacts for the i^{th} trade can be computed following Eq. (2) as $\widehat{InfPI_i} = \Delta p_{non-tr,i}$ and $\widehat{InvPI_i} = \Delta p_{tr,i} - \Delta p_{non-tr,i}$. As implied by Eq. (9-10) final estimates of the two price impacts are computed as average over all trades/iterations $InfPI_{Discr} = E(\widehat{InfPI_i})$ and can be then compared to the true value $InfPI$.

These simulations show that price discreetness does not affect the estimated price impact components if the number of trades is sufficiently large. I simulate the system for different parameter values, and the estimates of price impacts computed from simulated discrete prices always match perfectly their true values. The intuition is that price discreetness simply adds another source of noise which is averaged out in a sufficiently large sample.

However, price discreetness of course affects the distribution of individual price impacts by introducing significant skewness in it. If prices are continuous, information and inventory price impacts for individual trades have a symmetrical and continuous distribution (Panel A of Figure A1).[6] Thus, in this frictionless case, both the mean and the median produce the correct information and inventory impacts of 0.2 and 0.4 cents respectively. However, after a five-cent tick size is introduced in Panel B, individual price impacts can only take values from the (… -10,

---

[5] The information and inventory price impacts are set to $InfPI = 0.2$ and $InvPI = 0.4$ cents to match my sample of option trades. $\epsilon_{Inf,i} \sim N(0, 1.4^2)$, $\epsilon_{Inv,i} \sim N(0, 2.4^2)$, $Corr(\epsilon_{Inf}, \epsilon_{Inv}) = 0$.

[6] Importantly, public information error ϵ is assumed normal (symmetric) in Eq. (A3)

-5, 0, 5, 10 …) cent grid as implied by Eq. (A4). The average is unchanged and estimates the correct price impacts of 0.2 and 0.4 cents. But the median becomes zero because it must lie on the price grid. The distribution for the actual option trade data (Panel C) is surprisingly close to the simulated one (Panel B). Thus, price discreetness and my simulation procedure indeed capture key features of the options data. The two sets of distributions have the same means and variances, but the actual data has fatter tails perhaps because public information shocks have fatter tails than is implied by the normal distribution used in simulations. Panel C is based on a part of my main sample, where exactly two exchanges quote the trade price, and the expected changes in price are set to zero.

The median over individual information impacts is negative (-0.03%) in summary statistics in Table 1 because positive expected changes (0.03%) are subtracted from the median price response of non-trading exchanges, which is zero. The distribution of expected changes is continuous and approximately symmetric in the data.

Overall, as for most other microstructure methods, price discreetness does not introduce a bias in the price impact decomposition estimates.

A.4. Trade sign classification

The data do not specify whether a given option trade is initiated by a buyer or a seller, as a result trade direction is inferred from a comparison of trade price with quoted prices. The potential concern is that trade sign is misestimated for some trades.

This concern has been extensively studied in the microstructure literature because many popular microstructure methods (such as PIN) rely on trade direction, and the direction is not reported in standard databases such as TAQ. Odders-White (2000) among many others shows that standard methods such as Lee-Ready and the quote rule classify correctly about 85% of stock trades. The results are similar for the options market: Savickas and Wilson (2003) show that the quote rule signs correctly 83% of option trades. As a result, the sign misclassification usually introduces a modest downward bias in point estimates (similar to other types of estimation error) but does not affect main conclusions. For example, Boehmer et al. (2007) show that if all trades were classified correctly stock PIN would be 18% higher (i.e., 17.6% instead of 13.6%); however, this correction does not alter any conclusions in the original PIN paper.

Although the literature suggests that the error in estimating trade sign is usually not important, this error still may affect the price impact decomposition. By design, the decomposition is applied to a subsample of all option trades for which misclassifying trade direction is highly unlikely. Specifically, in my main sample, a trade is classified as a buy if (i) its price equals to the national best ask price and (ii) this best ask is quoted by at least two exchanges including the one that reports the trade. It is hard to imagine a scenario where a trade that satisfies these two conditions is a sell instead of a buy. For example, such a seller-initiated trade would violate the price-time priory enforced at most exchanges. Specifically, this sell trade is executed ahead of at least one sell limit order at the same exchange that offers the same price but has been submitted prior to it (also, at least one sell limit orders at other exchanges offers the same price). This is a clear violation of the time priority (the price is the same). Supporting the idea that misclassification is rare for such trades, Odders-White (2000) and Ellis et al. (2000) show that most of the 15% misclassification error in the stock market comes from trades executed inside the quotes. Indeed, it is hard to say whether a trade executed at the quote midpoint is a buy or a sell. Thus, for trades executed at either best bid or ask, the classification error is in the single digits. Extrapolating these results to my sample where at least two exchanges quote the trade price – the misclassification rate must be even lower.

Finally, I quantify the effect of trade sign misclassification on the inventory-risk and asymmetric-information components within the theoretical framework of Section 3.2. In short, the ratio of the two components is unaffected by the misclassification, but the estimates will underestimate the true magnitude by about twice the classification error. E.g., if the sign is misclassified for 10% of trades, the estimated inventory impact will be 20% lower than the true value. Thus if anything, trade misclassification makes it harder to find significant price impacts. To show this result, consider first an extreme case where the direction of all trades is misclassified, then the estimated inventory-risk impact will be simply the opposite of the true value: $\widehat{InvPI} = -InvPI$. This is implied by Eq. (10) where the difference in price responses is multiplied by buy-sell indicator – the difference is correct but the buy-sell indicator is wrong (e.g., -1 instead of 1 for buys).[7] More generally, if trade direction is misclassified for 10% of trades, then considering separately the subsamples of correctly classified and misclassified trades, the estimated price impact will be $\widehat{InvPI} = 0.9 * InvPI - 0.1 * InvPI = 0.8 * InvPI$,

[7] Equations (4) and (4') show that bid and ask prices move together. Equation (9) $\theta \cdot X = E\left[I_i^{BS} \cdot (\Delta p_{t,i-}^{BS} - E(\Delta\mu_t | F_t))\right]$ then implies that if a buy is misclassified as a sell, i.e., $I_i^{BS} = -1$ instead of 1, then information impact is $-\theta \cdot X$ is instead of $\theta \cdot X$.

the true impact is indeed underestimated by twice the error rate. Eq. (9) implies a similar result for the information price impact: $\widehat{InfPI} = 0.9 * InfPI - 0.1 * InfPI = 0.8 * InfPI$. Hence, the ratio between the two impacts is not affected by the misclassification error.

Overall, the method is designed to minimize the error in classifying trade direction. It is unlikely to affect paper's conclusions and if anything would strengthen them.

A.5. Is there "hot potato" trading in options?

This section confirms empirically that option market-makers do not commonly share inventory after a client trade; thus, validating one of the method's assumptions. The method assumes that market-makers do not regularly share inventory positions directly with each other. If they do, the price response from the non-trading market-makers cannot be attributed only to asymmetric information because they not only learn about the trade but also anticipate getting a chunk of it from the trading market-maker. In this case, the method will overstate the asymmetric-information impact and underestimate the inventory-risk impact. For example, if there are only two market-makers and the trading market-maker shares half of the trade size with the non-trading market-maker immediately after a trade; then Eq. (A5) and (A6) imply the identical price response for both market-makers. Thus, the method implies zero instead of positive inventory impact in this example.

$$\Delta p_{t,i}^{a} = E(\Delta\mu_t|F_t) + \theta\big(X - E(x_t|F_t)\big) + \gamma \cdot X + \epsilon_{t+\Delta} + (\vartheta_{t+\Delta t}^{a} - \vartheta_{t}^{a}) \qquad \text{(A5)}$$

$$\Delta p_{t,i-}^{a} = E(\Delta\mu_t|F_t) + \theta\big(X - E(x_t|F_t)\big) + \gamma \cdot 0 + \epsilon_{t+\Delta} + (\vartheta_{t+\Delta t}^{a} - \vartheta_{t}^{a}), \qquad \text{(A6)}$$

Market-makers may want to share large trades to reduce inventory risk. Ho and Stoll (1983) show this theoretically, while Reiss and Werner (1998) as well as Lyons (1996) report empirical evidence from the equity and foreign exchange markets based on data from early nineties. More recently, the role of dealers is diminishing in both markets, so interdealer trading is less prevalent now. In the options market, market makers are the main liquidity providers, so hot potato trading could potentially be important there.[8]

If hot potato trading is common in options, then (by analogy with Lyons, 1996)

[8] For example, Jameson and Wilhelm (1992) based on "casual observation of trading activity on the exchange floor" claim that sharing of the incoming orders was common among CBOE market-makers in mid-80s. However, they don't provide any further evidence for the claim.

immediately[9] after each client transaction, market-makers will initiate a sequence of back-to-back smaller trades in the same option contract to redistribute the incoming trade. Moreover, inventory sharing should be more prevalent after large trades. However, the trade sequences can be produced for many other reasons, for this reason hot potato trading is hard to identify empirically. For example, investors may split a large order into smaller pieces, or they can respond to the same news.

Thus, if the options data contain a lot trade sequences, it may or may not indicate hot-potato trading; but if the trade sequences are rare even after large trades, then the inventory sharing between market makers is rare too. The data support the later alternative.

Table A1 reports the number of trades in the same option contract in one-minute interval around a trade for both the entire sample as well as for one percent of largest trades. Hot-potato trading would trigger a lot of trading activity especially for large orders. However, most trades (59%) have no other trades around them. Importantly, trade sequences are no more likely after the largest trades. Even for 9% of trades with more than five other trades near them (it will take five trades to fully share a trade between six exchanges in my sample), it's likely that most of them are not interdealer trades. Also, this number grossly overestimates the relative frequency of large trade sequences as each sequence is counted once for every trade it contains. Importantly, the sample contains a complete set of trades for a given option including all potential interdealer trades. According to exchange rules fiercely protected by the SEC, all option trades must be exposed to the public through option exchanges making it's hard for market-makers to internalize trades.

Overall, hot-potato trading is not common in the options market.

A.6. Are prices at all exchanges equal?

The price impact decomposition identifies the inventory impact by taking the difference between price responses of the trading and non-trading exchanges. Therefore, it is important to explain why price quotes from different market-makers and exchanges mean the same thing. To illustrate this concern, imagine that one exchange accounts for all the price discovery and trading, while other exchanges simply follow it. That is, other exchanges are a side show with little trading. This used to be the case in the equity market when NYSE dominated it while

[9] Major option market-makers are completely electronic since early 2000s, so "hot potato" trading must be wired in their computer algorithms.

regional exchanges were a side show.

Option prices from all exchanges mean the same thing because (like equities) all US exchange-traded equity options are centrally cleared. Option Clearing Corp. (OCC) is counterparty on all option trades since 1973. Second, during my sample period, all option exchanges had similar market structure (SEC Report, 2007) dominated by electronic trading and characterized by "payment for order flow." Obviously, the exchanges are not totally identical; for example, there are some differences in technology and fees. Third, a dozen of market-makers dominate the option liquidity provision and use similar algorithms.[10] For each option class, each exchange assigns a different lead market-maker. For example, Citadel makes the market in Google options at ISE while Susquehanna does this at CBOE and vice versa for options on Yahoo.[11] Competition is high between the option exchanges: SEC (2007) reports that for the entire option universe at least four exchanges quote the best bid price during about 78% of a trading day.[12] Finally, given these observations it is not surprising that all option exchanges participate in option price discovery: Simaan and Wu (2007) show that the Hasbrouck information share ranges from 8% for PHLX to 17.9% for ISE in January 2002.

For robustness, I confirm that the main results are not driven by trades from a single options exchange. First, market share is not concentrated at a single exchange: about one third of all trades in the sample are executed at ISE (34%), followed by CBOE (27%), PHLX (12%), and Pacific (11%) as reported in Table A2. These numbers match closely the overall market shares of option exchanges based on the entire equity options universe. Second, I compute the asymmetric-information and inventory-risk components for subsamples where trades from a given option exchange are excluded from the main sample. The information impact ranges from 0.17% to 0.23%, while the inventory-risk impact varies from 0.38% to 0.45%.

Overall, although price impacts vary by exchange, this variation does not alter the main conclusions of the paper.

A.7. Outliers

Although my sample size is large, outliers can still present a problem if there are

[10] For example, in 2004, the list of lead option market-makers include Citadel, Citigroup, Credit Suisse, Deutsche Bank, Knight, Morgan Stanley, SLK-Hull (later became part of Goldman Sachs), Susquehanna, Timber Hill, UBS and Wolverine Trading.

[11] For example, in its letter to SEC in 2005, Citadel Derivatives describes itself as "an options market maker, active on all six options exchanges, including acting as a specialist on the ISE, CBOE and PCX."

[12] http://www.sec.gov/news/studies/2007/optionsroutingreport.pdf

sufficiently many of them. This section shows that outliers have a negligible effect on the estimates of the information and inventory components reported in the paper. Also, the most extreme outliers have been already removed from the main sample. As discussed in Section 4.2, a trade is removed from the main sample if the absolute value of at least one of the price impact components is greater than 50% for it. Table A3 reports the estimates for information and inventory components removing outliers based on the different values for this threshold. The threshold ranges from the case of no threshold down to the threshold of 30%. Without the threshold no observations are removed, and the information and inventory components are 0.217% and 0.446% respectively. For the threshold of 80%, 152 observations are dropped and the price impacts decrease to 0.216% and 0.414%. If the threshold is lowered to 30%, 4861 trades are dropped and the components decrease to 0.212% and 0.40%. The estimates decrease slightly because some of the dropped observations were false positive (not outliers) particularly for lower threshold values, some trades do have a large price impact.

Overall, outliers have an insignificant impact on the price impact components in my sample.

A.8. Information impact for stocks vs. ETFs

The price impact decomposition is similar to other methods in that it estimates the information content of trades but is silent about what kind of information stands behind it. However, a comparison between the information price impacts for ETFs and individual stocks may shed some light on this important question. The sample includes four ETFs on stock indices, which account for about one-fifth of all option trades. As can be inferred from Table 2, a trade in option on ETF has information impact of 0.15%, which is below the impact for individual stocks (0.23%); and inventory impact of 0.40%, which is similar to stocks (0.42%). Inventory risk works similarly for ETFs and individual stocks. However, ETFs and stocks may differ in the type of private information, particularly for options. Consider first the equity market, prices and weights of ETF constituents are known and the creation/redemption mechanism insures that ETF price stays close to its value. Since getting private industry-wide information is perhaps harder than private stock-specific information, less informed trading is expected in ETFs. Consistent with this idea, Table 2 shows that option trades have large impact on the underlying price for stocks but almost no impact for ETFs (0.022 vs. 0.003 per trade). I.e., option trades contain little new information about the ETF price level.

However, moving to the options market, options on ETF and options on its constituents are not linked as tightly. ETF option is an option on a basket of stocks. Thus, ETF option price depends on (i) constituents' prices, (ii) their volatilities (can be inferred from their options), (iii) the correlation between constituents' prices (historical estimates are noisy) and finally (iv) an option pricing model that aggregates these components. Thus, trades in ETF options could be informed because some investors have superior information about the model inputs (such as the correlation), or simply have a better option pricing model. According to this intuition, reasonable amount of informed trading in ETF options is expected but perhaps less than in individual stock options, which is what I found in my sample.

Finally, my conclusions here should be taken with caution as the number of ETFs in the sample is small and the argument is based on general considerations, not a rigorous model.

A.9. Alternative channels for return predictability in instrumental variables approach

What is a channel through which past order flow predicts future returns? This section shows that order imbalance predicts future returns predominantly through its ability to predict future order imbalances. Thus, most of the predicted order imbalances can be attributed to future inventory shocks.

Order imbalance can affect option returns through two alternative channels. Chordia and Subrahmanyam (2004) advocate the inventory-risk channel. Their model implies that the order imbalance on day *t-1* helps to predict the order imbalance on the next day *t* which in turn moves option prices on the same day *t*. The alternative channel is informed trading advocated by Ni, Pan and Poteshman (2008). They find that option traders have private information about future stock volatility. Thus, order imbalance can predict future stock volatility which in turn directly transmits into future option returns. However, unlike data in Ni et al. (2008), the ISE order flow data are public and thus create little potential for informed trading.

I adjust an instrumental variables approach to compare the two channels and show that the inventory-risk channel dominates the informed-trading channel. In the first stage of 2-SLS, current order imbalances and volatility are instrumented with their past values. In the second stage, day *t* option returns are regressed on the predicted day *t* order imbalances and volatility. If the inventory channel dominates the information channel, then the instrumented volatility will have small or no predictive power on returns.

I choose the same measures of day *t* order imbalance as before: market-wide and individual imbalances. Stock volatility is measured in two complementary ways. The first measure is absolute stock return. However, volatility may vary widely across stocks and in time (ARCH effects). The second measure tries to account for these two features. An adjusted absolute return is computed as an absolute stock return normalized by its 20-day moving average. It measures how high current volatility is relative to the recent past. In addition, regression coefficients for this measure are easier to interpret. Overall, the two measures complement each other. I use six instruments: lag of market-wide order imbalance, two lags of individual imbalance, lag of order imbalance for short-term options as well as lag of two volatility measures: absolute stock return and the adjusted absolute return.

The first half of Table A6 reports the first stage for the 2-SLS regressions. As have been discussed already, current and future order imbalances are positively correlated. Next, I confirm the results of Ni et al. (2008) that order imbalances predict future stock volatility. However, this predictability is not necessary driven by informed trading. In particular, market-wide imbalance is the most significant predictor of future volatility, but market-wide variables are unlikely to be affected by informed trading. It is hard to obtain private information about the entire market. The alternative explanation for the order flow predictability of volatility is that the econometric model does not account for information about future volatility that the market already knows. Investors hedge or speculate by buying options and thus, create order imbalance before an expected future event which causes a volatility spike.[13]

The last four columns of Table A6 report the second stage of 2-SLS. The regression in Column 6 estimates the sensitivity of option returns to the same-day order imbalance and a coefficient of 0.078 means that the order imbalance of 25%, which equals to one standard deviation, corresponds to option returns of about 2%. The next column adds the market-wide order imbalance and the coefficient drops by half indicating that the market-wide imbalance is at least as important as individual imbalance.

In the last two columns, I conduct a horse race between the inventory-risk and volatility channels for the return predictability. Both future volatility and order imbalance are instrumented

[13] For example, economic releases such as GDP are often associated with high volatility. The timing of the releases is known many days in advance to all market participants but is not known to the econometrician who relies only on history of prices and volumes. In anticipation of the release, some investors adjust their portfolio to hedge or speculate. Usually, such trades are correlated and they create order imbalance. This imbalance is observed by the econometrician who concludes that the imbalance predicts future volatility. However, in this example, there is no informed trading since information is common to all investors.

and placed in the same regression. Both measures of volatility are insignificant if measures of order imbalance are included. At the same time, the coefficient estimates for order imbalances are unchanged. This result indicates that past order imbalance predicts future returns through future order imbalance rather than through future stock volatility. It provides further support for the important role that inventory risk plays in option pricing.

Another concern is that individual order imbalances contain some private option-specific information that becomes known to the market and realized in option returns only on subsequent days. The effect of this channel is likely to be small for two reasons. First, day *t-1* individual order imbalance becomes public information by the end of that same day. Subscribers to the ISE open/close data receive updated estimates of the order imbalance every ten minutes.[14] Even without special data products from ISE and CBOE, the order imbalance can be estimated from the public tick-level data broadcasted by OPRA in real time. Second, the effect of private information embedded in the lagged order imbalance should remain significant even after controlling for future order imbalances. However, in untabulated results, I show that this is not the case. The coefficient for day *t-1* individual order imbalance decreases from 0.008 to 0.002 and becomes economically insignificant[15] if day *t* order imbalance is included in the return regression.[16] This result indicates that individual imbalance predicts future returns predominantly through future individual imbalance and thus, is unlikely to be driven by private information. This result also confirms one of the main tests suggested by the Chordia and Subrahmanyam (2004) model. The result is consistent with Barber et al. (2009) who find that the trading of retail stock investors is highly correlated and persistent.

A.10. Other predictors of future option returns

Although the paper's primary attention is on the link between returns and order flow, this section discusses other significant predictors of option returns. These variables generally have smaller predictive ability than past order imbalances.

The battery of controls includes about fifty variables, many of which have not previously been considered in the literature. Table A7 reports the most significant control variables in the regression from Eq. (16). Column 3 reports coefficient estimates for the whole sample, while other columns examine particular subsamples. Column 4 reports the results for the subsample of

[14] See http://www.ise.com/market-data/products/put-call-data/ise-open-close-trade-profile-intraday/ .
[15] A coefficient of 0.002 times standard deviation of 0.3 corresponds to a 0.06% changes in option returns.
[16] The regression option returns on day *t* on explanatory variables and order imbalances from day *t-1*.

two hundred stocks with most liquid options. Column 5 studies a subsample of options with call and put bid prices exceeding 2 dollars. Column 6 reports results for two-day-ahead returns. These three subsamples aim to examine how market microstructure influences the return predictability. Finally, the last column reports univariate regressions of option returns on a single variable and an intercept.

$$\text{OptRet}_{t,i} = \alpha_0 + \alpha_1 \text{OrdImb}_{t-1,i} + \alpha_2 \text{MWOrdImb}_{t-1} + \beta' \text{OtherPredictors}_{t-1,i} + \varepsilon_{t,i} \qquad (16)$$

Even among the variables reported in the table, few are economically significant and stable across model specifications. In addition, several variables are significant if all controls are included but become insignificant in univariate regressions.

Absolute stock return is a good predictor of next-day option returns. One standard deviation change in this variable increases returns by 0.39%.[17] That is, option prices "underreact" to changes in instantaneous volatility. Poteshman (2001) finds a similar result for S&P index options. He documents that S&P index options underreact to unexpected changes in instantaneous volatility estimated from a stochastic volatility model. Specifically, he estimates a regression of a difference between changes in instantaneous volatility for long-term and short-term options on unexpected changes in instantaneous volatility. The coefficient is negative but insignificant. Poteshman interprets his findings as evidence of investor irrationality. Investors put too much weight on the prior beliefs and do not update them properly. My paper differs from Poteshman (2001) in several ways. My methodology is less sophisticated, as returns are computed directly instead of relying on a specific model. Consequently, the economic magnitude is easier to estimate with this approach. Finally, I study equity options while Poteshman examines S&P index options.

Although it is tempting to blame investor irrationality for this return predictability, I suggest an alternative explanation. A large tick size in the options market may be the reason why option prices are "sticky" and unresponsive to small changes in volatility. To support this microstructure hypothesis in unreported results, I show that the coefficient for absolute returns becomes virtually zero after November 1, 2009 (and February 2010). At that time, the majority of stocks were added to the penny pilot program that reduced the options tick size from 5 cents to a penny. At the same time, the coefficient for the previous year starting on November 1, 2008 is as big as for the full sample. This difference-in-difference result favors the microstructure

[17] A related factor is median stock volume which is significant only when absolute return is omitted.

explanation. In addition, the predictive ability of absolute return is much smaller for subsamples of liquid options and options with large price.

Another variable with significant predictive power is one-day change in ATM implied volatility. If implied volatility increases by one standard deviation (4%), option returns become lower by 0.39% on the next day. The "bid-ask midpoint bounce" can explain this predictability. To illustrate the mechanism, consider a stylized example. If the option bid price is set abnormally low, the quote midpoint will also be low which translates into low implied volatility. The bid price and implied volatility revert to the normal levels on the next day; and positive option returns are recorded on this day because returns are computed based on the quote midpoints. However, no abnormal returns will be recorded if the ask price is used instead of the midpoint in this example. Thus, the decrease in implied volatility on day *t-1* is reversed on day *t* and is mechanically related to the option returns on day *t*. The microstructure explanation is supported by the fact that the predictability is much smaller in the subsample of the two hundred most liquid stocks. Finally, there is no predictability between the change in implied volatility on day *t-1* and the option returns on day *t+1* which is directly implied by the "midpoint bounce" explanation.

Everything said about the change in implied volatility applies to the implied volatility as a predictor. It is also likely driven by microstructure reasons because its predictive ability disappears if day *t+1* returns instead of day *t* returns are predicted. Predictive ability of lagged option returns is also mainly driven by the market microstructure. Surprisingly, it is not significant in the individual regression.

Jones and Shemesh (2010) show that option returns are abnormally negative over the weekend (Friday to Monday close). Confirming their findings, I also find that the weekend returns are lower by 1.3% than on weekdays, which is somewhat lower in magnitude than a -1.8% return found by Jones and Shemesh. The half percentage point difference can be partially explained by the difference in sample periods. Jones and Shemesh use data from 1996 to 2007 and find in their Table 9 that the weekend effect becomes much weaker towards the end of this period; while my sample period starts in 2005. The economic mechanism for the weekend effect is not clear; however, it can be partially driven by microstructure effects. First, option bid-ask spread decreased sharply after the launch of ISE in 2001, so did the weekend effect. Second, the weekend effect is much weaker for the subset of options with price above 2 dollars and completely disappears if the expiration weekend is excluded from this option set. Also surprisingly, the weekend effect is much smaller outside of the expiration weekend in my sample but not in theirs. In untabulated results, I show that except for the expiration period, order imbalance exhibits little day-of-week seasonality. Thus, order flow is not responsible for day-of-week seasonality in option returns. Overall, the weekend effect is a particularly intriguing market anomaly that requires more academic research to uncover its economic mechanism.

Boes et al. (2007) show that close-to-open jump risk is priced for S&P index options. The idea is that stocks for which most of the volatility happens during non-trading hours[18] should have more negative option returns to compensate investors who are short gamma for inability to hedge during non-trading hours. I find similar risk premium for equity options. Stocks with higher close-to-open volatility relative to close-to-close volatility produce more negative option returns in the future. The economic magnitude varies from -0.1% per day for the full sample to -0.3% for the subsample of the two hundred most liquid stocks.

The fact that my results are generally consistent with other theories reported in the literature indicates that my specification is reasonable. Overall, the examination of other explanatory variables is consistent with the conclusion about order imbalance being a major predictor of future option returns.

A.11. Robustness tests for return predictability

This section confirms that the result that past order flow predicts future returns is robust to changes in methodology and different subsamples. The order imbalance from day *t-1* predicts option returns on the next day *t*, but can it predict returns two days ahead on day *t+1*? Column 3 of Table A8 shows that it indeed can. Moreover, the coefficients are very similar to the baseline case reported in Column 2. Column 4 reports results for the day *t+4* returns (one week ahead). For this case, the market-wide imbalance remains highly significant, while the individual imbalance has smaller magnitude. One standard deviation change in the predicted order imbalance corresponds to a 0.6% return on day *t+1* and a 0.5% return on day *t+4*, which is smaller than 1% for day *t* but is still large. Thus, the returns are highly predictable for several days in the future. The results are robust to the way option returns are computed. Column 5 uses delta-neutral call returns instead of delta-neutral straddle returns and finds little change in

[18] Specifically, I look at the difference between close-to-open volatility and time-scaled standard volatility. The time scaling is done to make the mean of the variable approximately zero.

coefficient estimates. My sample includes all equity options with at least some trading activity listed on ISE with a maximum of 1911 stocks on a given day. Many of these stocks have illiquid options with only few trades per day. A potential concern is that stocks with illiquid options drive the return predictability. I test this hypothesis on the subsample of the 200 stocks with most liquid options based on volume in the previous 250 days. Column 6 shows that the predictability for stocks with liquid options is very close to the baseline case.

The last column in Table A8 reports the most important robustness check. Inventory risk is higher during periods of market stress such as financial crises as market-makers are more risk-averse, and markets are more volatile. Indeed, order imbalance has a higher impact on option prices during crises. I study the interaction between order imbalance and the crisis dummy variable which is set to one for the period August 2007–January 2009. The results are similar if VIX is employed instead of the dummy variable. The last column in Table A8 reports that the crisis dummy is mechanically positively related to option returns because market volatility increases during the crisis leading to positive straddle returns. The main coefficients of interest are interaction terms between the crisis dummy with market-wide and individual order imbalances. Both coefficients are highly economically significant. For example, market-wide order imbalance has almost two times bigger price impact during the crisis compared to normal time.

B.1. Variable description

Name	Description	Computation
Option returns		
OptRet(p=0)	Straddle returns for expiring options (T-t < 13)	
OptRet, OptRet(p=1)	Straddle returns for short-term options	See equation (11)
OptRetCall(p=1)	Call option returns for short-term options	See equation (11)
OptRet(p=3)	Straddle returns for long-term options	See equation (11)
Order imbalance		
OrdImb	Order imbalance	See equation (12)
OrdImb(p=1)	Order imbalance for short-term options	See equation (12)
$OrdImb_{t-1}$	Order imbalance on the previous day	See equation (12)
OrdImbPut	Order imbalance for put options	See equation (12)
MWOrdImb	Market-wide order imbalance	See equation (13)
AdjOrdImb	Volume-adjusted order imbalance	See equation (12)
Dummy variables		
n_1	Expiration day (Friday)	1, if t = expiration
n0	Post-expiration day (Monday)	1, if t-1 = expiration
nead0	Pre-earnings announcement day (pre-EAD)	1, if t+1 = EAD
nead1	Earnings announcement day (EAD)	1, if t = EAD
Weekend	Weekend dummy	1, if t = Friday
n_crises	Crisis dummy	1, if August 1, 2007 < date < Febuary 1, 2009
Stock market		
StkRet	Stock returns	
AbsStkRet	Absolute stock returns	|StkRet|
RelAbsStkRet	Absolute stock returns normalized by its average over the previous 50 days	
$Log(Open/Close_{t-1})$	Close-to-open ratio	$\log(Open_t / Close_{t-1})$
Log(High/Low)	High-low ratio	$\log(High_t / Low_t)$
RelLog(High/Low)	High-low ratio normalized by its average over the previous 50 days	
StkPrice	Close stock price	
logME	Logarithm of market capitalization	
StkVol	Stock volume in dollars	
MeanStkVolume	Stock volume in dollars, a 75-day moving median	
Momentum	Stock price relative to its 250-day moving median	$\frac{\log(Close/MedianClose)}{StdRet * \sqrt{250}}$
StkRet1W	One-week stock returns	
StkRet1M	One-month stock returns	
StkRet6M	Six-month stock returns	

Stock volatility		
StdRet	Stock returns volatility (based on a 50-day moving average)	Mean{ log($Close_t$ / $Close_{t-1}$)^2 }
COStdRet	Close-to-open volatility (based on a 50-day moving average)	$\sqrt{Mean\{\log(Open_t / Close_{t-1})^2\}}$
HLStdRet	High-low volatility (based on a 50-day moving average)	$\sqrt{1/4\log 2}\sqrt{Mean\{\log(High/Low)^2\}}$
AbsStdRet	Absolute returns volatility (based on a 50-day moving average)	$\sqrt{\pi/2}\sqrt{Mean(AbsStkRet)}$
Implied volatility/options market		
IV30	ATM implied volatility for 30-day options	
IV60	ATM implied volatility for 60-day options	
IV60d25	Implied volatility for 60-day put options with delta of -0.25	
IV360	ATM implied volatility for one-year options	
IV360d25	Implied volatility for one-year put options with delta of -0.25	
IV(p=1)	Average implied volatility for short-term options	
diff(IV)	One-day change in short-term implied volatility	IV(p=1) - IV(p=1)$_{t-1}$
IV30 - IV30$_{t-1}$	One-day change in implied volatility	IV30 - IV30$_{t-1}$
IV30 - IV30$_{t-5}$	One-week change in implied volatility	IV30 - IV30$_{t-5}$
IV30 - IV30$_{t-25}$	One-month change in implied volatility	IV30 - IV30$_{t-25}$
Skew60	Volatility skew for 60-day options	log(IV60d25/IV60)
RNVolatility	Risk neutral volatility as in Bakshi, Kapadia, and Madan (2003), 50 days to expiration	See Bakshi et al. (2003)
RNSkewness	Risk neutral skewness as in Bakshi, Kapadia, and Madan (2003), 50 days to expiration	See Bakshi et al. (2003)
RNKurtosis	Risk neutral kurtosis as in Bakshi, Kapadia, and Madan (2003), 50 days to expiration	See Bakshi et al. (2003)
IV60 - IV360	Volatility time slope defined as the difference between 60-day and 360-day implied volatilities	IV60 - IV360
IV60-StdRet	Volatility premium defined as the difference between 60-day implied volatility and stock volatility	IV60 - StdRet
COStdRet-StdRet	Close-to-open relative to close-to-close volatility	2*COStdRet - StdRet
VolCone(IV30)	Volatility cone, 30-day IV relative to its average over previous year	(IV30-Median(IV30, 250-days))/ (Range(IV30,250-days)/1.349)
IdioIV30	Idiosyncratic volatility proxy	30-dayIV^2 - VIX^2
VIX	CBOE VIX (S&P500 30-day implied volatility)	
OptBidAsk	Average bid-ask spread for ATM short-term options	
Option volume		
OptVolumeUSD	Option volume in dollars	
OptVolume	Option volume measured in contracts	
MeanOIUSD	Logarithm of open interest, a 250-day moving average	
MeanOptVolumeUSD	Logarithm of option dollar volume, a 250-day moving average	

RelVolUsd (p=1)	Dollar volume for short-term options relative to its 520-day median	Log(OptVolUsd(p=1)) - Median(LogOptVolUsd(p=1))
RelOICall(p=1)	Open interest for short-term call options relative to total short-term open interest	OICall(p=1)/OI(p=1)
RelOICallITM(p=1)	Open interest for short-term ITM call options relative to total short-term open interest	OICallITM(p=1)/OI(p=1)
RelOIPutOTM(p=1)	Open interest for short-term OTM put options relative to total short-term open interest	OIPutOTM(p=1)/OI(p=1)
RelVolCall(p=1)	Volume for the short-term call options relative to total short-term volume	OptVolCall(p=1)/OptVol(p=1)
RelVolCallITM(p=1)	Volume for the short-term ITM call options relative to total short-term volume	OptVolCall(p=1)/OptVol(p=1)
RelVolPutOTM (p=1)	Volume for the short-term OTM put options relative to total short-term volume	OptVolCall(p=1)/OptVol(p=1)
Vol/OI(p=1)	Volume relative to open interest for short-term options	OptVol(p=1)/OI(p=1)

Reference

Boehmer, Ekkehart, Joachim Grammig, and Erik Theissen. "Estimating the probability of informed trading—does trade misclassification matter?." Journal of Financial Markets 10.1 (2007): 26-47.

Chordia, Tarun, and Avanidhar Subrahmanyam. "Market making, the tick size, and payment-for-order flow: theory and evidence." Journal of Business (1995): 543-575.

Dravid, A.R., 1991. Effects of bid-ask spreads and price discreteness on stock returns. Working Paper. Rodney L. White Center, Wharton School.

Ellis, Katrina, Roni Michaely, and Maureen O'Hara. "The accuracy of trade classification rules: evidence from Nasdaq." Journal of Financial and Quantitative Analysis 35.04 (2000): 529-551.

Glosten, Lawrence R. "Is the electronic open limit order book inevitable?." The Journal of Finance 49.4 (1994): 1127-1161.

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

Hasbrouck, Joel. "Measuring the information content of stock trades." The Journal of Finance 46.1 (1991): 179-207.

Hasbrouck, Joel. "The dynamics of discrete bid and ask quotes." The Journal of finance 54.6 (1999): 2109-2142.

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.

Kandel, Eugene, and Leslie M. Marx. "Nasdaq market structure and spread patterns." Journal of Financial Economics 45.1 (1997): 61-89.

Kyle, Albert S. "Informed speculation with imperfect competition." The Review of Economic Studies 56.3 (1989): 317-355.

Lyons, Richard K. "Foreign exchange volume: sound and fury signifying nothing? The microstructure of foreign exchange markets." University of Chicago Press, 1996. 183-208.

Muravyev, Dmitriy, Neil D. Pearson, and John Paul Broussard, "Is there price discovery in equity options?", Journal of Financial Economics 107.2 (2013): 259-283.

Muravyev, Dmitriy, and Neil D. Pearson (2014) "Option Trading Costs Are Lower Than You Think" Working paper.

Ni, Sophie X., Jun Pan, and Allen M. Poteshman. "Volatility information trading in the option market." The Journal of Finance 63.3 (2008): 1059-1091.

Odders-White, Elizabeth R. "On the occurrence and consequences of inaccurate trade classification." Journal of Financial Markets 3.3 (2000): 259-286.

Poteshman, A. M. 2001. Underreaction, Overreaction, and Increasing Misreaction to Information in the Options Market. The Journal of Finance 56, 851-876.

Reiss, Peter C., and Ingrid M. Werner. "Does risk sharing motivate interdealer trading?." The Journal of Finance 53.5 (1998): 1657-1703.

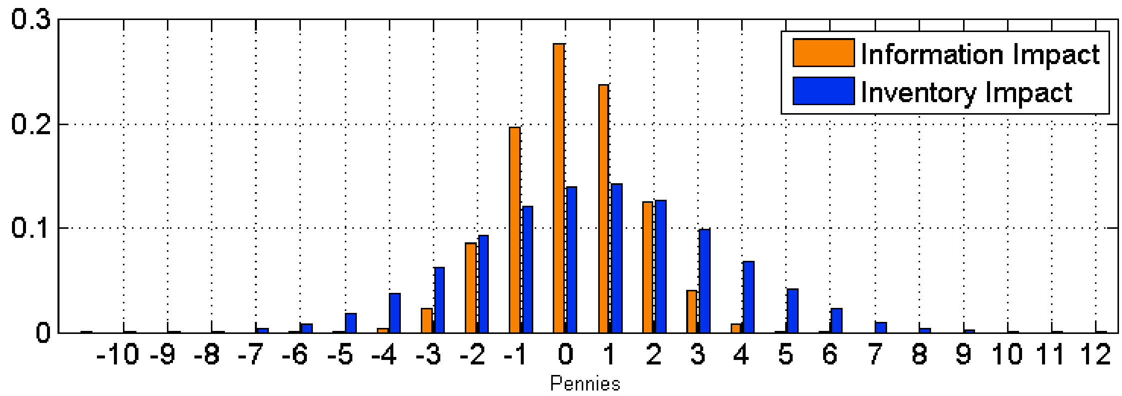
Savickas, Robert, and Arthur J. Wilson. "On inferring the direction of option trades." Journal of Financial and Quantitative Analysis 38.04 (2003): 881-902.

Securities and Exchange Commission. "Report Concerning Examinations of Options Order Routing and Execution" (2007).

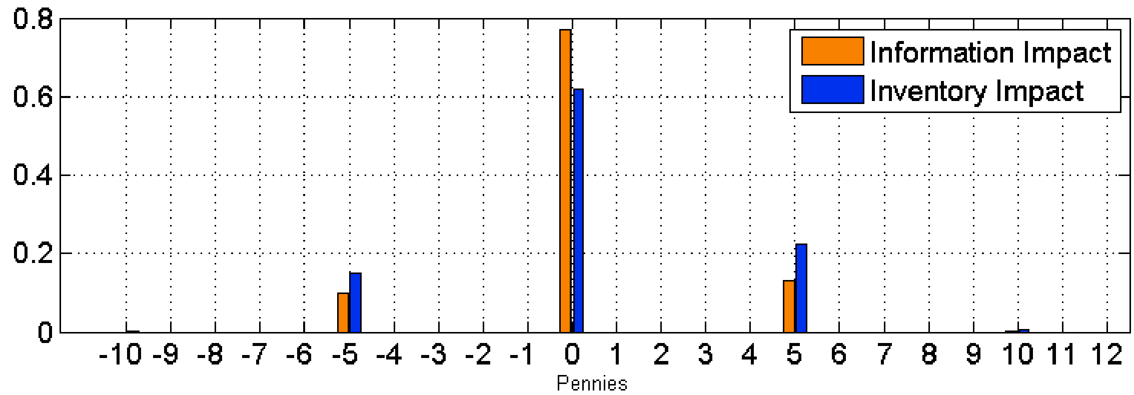
Simaan, Yusif E., and Liuren Wu. "Price discovery in the US stock options market." The Journal of Derivatives 15.2 (2007): 20-38.

Figure A1. Price discreetness and price impact components. The effect of price discreetness on the price impact components is simulated in Panels A and B and then compared to actual data in Panel C. Each panel shows a distribution for each of the two price impact components, based on price responses to individual trades computed following Eq. (A3) and (A4). Panel A considers frictionless case with zero tick size. Panel B sets tick size to five cents for the same sample of trades. Panel C show the distribution for the subsample of option trades with one trading and non-trading exchanges quoting the trade price.

Panel A. Zero tick size/continuous prices



Panel B. Five-cent tick size



Panel C. Actual data (option trades)

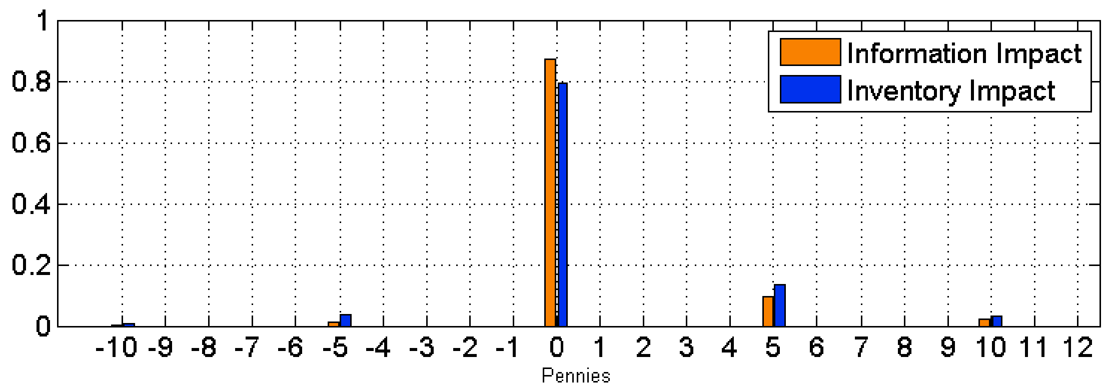


Table A1. Hot potato trading. For each trade, I compute the number of other trades in the same option contract that are close in time to the trade. The frequency for each number of trades is reported. For example, 16% of trades have only one other trade nearby. The frequencies are computed for the full sample as well as for one percent of largest trades for two time windows - one minute around and one minute after a trade.

	number of trades in -30 to 30 second window						
	0	1	2	3	4	5	>5
All trades	59%	16%	7%	4%	3%	2%	9%
Largest trades (1%)	56%	18%	8%	5%	3%	2%	8%

	number of trades in 0 to 60 second window						
	0	1	2	3	4	5	>5
All trades	63%	15%	7%	4%	2%	2%	7%
Largest trades (1%)	61%	16%	7%	4%	3%	2%	6%

Table A2. Price impacts by trading exchange. The table reports percentage of option trades executed by each option exchange. It also reports the asymmetric-information and inventory-risk components of price impact if all the trades from a given exchange are excluded from the main sample. BOX lunched in February 2004.

Exchange	Share of Trades	Asymmetric Information, %	Inventory Risk, %
ISE	34.2%	0.21	0.41
CBOE	26.6%	0.23	0.45
PHLX	12.4%	0.22	0.45
Pacific	10.9%	0.23	0.38
BOX*	9.9%	0.17	0.38
AMEX	6.0%	0.23	0.41

Table A3. Outliers. The table reports the number of trade observations removed for each threshold value and the estimates of the asymmetric-information and inventory-risk price impacts after the outliers are removed. A trade is classified as an outlier if the absolute value of at least one of the two price impact components exceeds the threshold (e.g., 50%).

Threshold, %	# of Outliers Removed	Information Impact, %	Inventory Risk Impact, %
30	4,861	0.212	0.400
40	2,197	0.215	0.407
50	498	0.216	0.413
60	298	0.216	0.413
70	193	0.216	0.413
80	152	0.216	0.414
n/a	0	0.217	0.446

Table A4. Expected changes in the option bid (ask) price. I estimate a regression of five-second changes in option bid (ask) quote on the lagged changes in option bid (ask) price and delta-adjusted stock quote midpoint.

$$\Delta p_{t,i}^{a} = \alpha_0 + \alpha_1\left(p_{t,i}^{a} - \mu_t^{BBO}\right) + \sum_{n=1}^{12} \alpha_{n+1} \Delta p_{t-n\Delta t}^{a} + \sum_{n=1}^{12} \alpha_{n+13}(\Delta_{t-n\Delta t} \cdot \Delta S_{t-n\Delta t}) + \epsilon_{t,i}$$

where $\left(\mu_t^{BBO} - p_{t,i}\right)$ is average quote midpoint across all exchanges minus the current exchange bid (ask) price. The lagged quote changes are based on twelve regularly spaced five-second time intervals (only the first two and the last coefficient are reported). The regression is estimated separately for bid and ask prices for each stock and six absolute delta (0.4 and 0.6 cut-offs) and time-to-expiration (70 days cut-off) bins within each day; average coefficients and R^2 are reported. Only observations with quoted price equal to NBBO ask at time t are included. All quote changes are measured in cents. All the coefficients are statistically significant with a minimum t-statistic of twelve.

	Money-ness	Time-to-Expiration	Inter-cept	BBO Deviation	Changes in option quoted price				Stock price changes adjusted for option delta				R^2
					t-1	t-2	…	t-12	t-1	t-2	…	t-12	
Bid	OTM	short-term	-0.016	-0.010	-0.119	-0.076		-0.015	0.198	0.093		0.020	0.036
		long-term	-0.016	-0.009	-0.120	-0.076		-0.015	0.218	0.101		0.021	0.033
	ATM	short-term	-0.034	-0.022	-0.142	-0.096		-0.019	0.265	0.133		0.028	0.057
		long-term	-0.032	-0.017	-0.131	-0.085		-0.018	0.262	0.125		0.026	0.042
	ITM	short-term	-0.046	-0.029	-0.157	-0.107		-0.023	0.300	0.149		0.031	0.057
		long-term	-0.044	-0.023	-0.138	-0.092	…	-0.020	0.275	0.133	…	0.028	0.046
Ask	OTM	short-term	0.016	-0.010	-0.120	-0.076		-0.015	0.194	0.091		0.019	0.036
		long-term	0.017	-0.010	-0.121	-0.077		-0.015	0.219	0.102		0.021	0.034
	ATM	short-term	0.037	-0.025	-0.142	-0.095		-0.020	0.265	0.131		0.028	0.056
		long-term	0.033	-0.021	-0.132	-0.086		-0.018	0.264	0.126		0.026	0.042
	ITM	short-term	0.048	-0.032	-0.156	-0.107		-0.022	0.299	0.150		0.030	0.057
		long-term	0.046	-0.027	-0.145	-0.096	…	-0.020	0.283	0.138	…	0.027	0.047

Table A5 Average correlations for selected variables. Options are divided into four days-to-expiration groups. Ultra short-term ("p=0") with less than 13 calendar days to expiration; short-term ("p=1") with on average 30-days; mid-term ("p=2") with no more than 150 days; and long-term ("p=3") with more than 150 days. Option returns ("OptRet") are computed for a delta-neutral straddle portfolio (long) based on the call-put pair which is closest to at-the-money. Returns are reported for the current (t) and next days (t+1). The order imbalance is based on the difference between the number of buy and sell trades normalized by the total number of trades on a given day ("OrdImb"). "MWOrdImb" are market-wide order imbalances. "OptBidAsk" is the dollar option bid-ask spread for short-term options. "OptVolume" is option volume measured in contracts. "IV30" is 30-days-to-expiration implied volatility. "Skew60" is a logarithm ratio of OTM to ATM implied volatilities for 60-days-to-expiration options. "RNSkewness" is risk-neutral skewness. "Diff(IV)" is one day change in short-term implied volatility. "IV60 - StdRet" is the difference between implied and historical volatilities. Other variables are defined in Section B of this internet appendix.

1	$OptRet_{t+1}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	OrdImb	0.03																	
3	MWOrdImb	0.08	0.16																
4	AbsStkRet	0.03	0.04	0.07															
5	Weekend	-0.06	0.01	0.02	-0.04														
6	StkRet	0.00	-0.06	-0.22	-0.01	0.00													
7	$OptRet_t$	-0.02	0.12	0.18	0.48	-0.02	-0.10												
8	Diff(IV)	-0.04	0.12	0.19	0.04	-0.05	-0.34	0.41											
9	StkRet1Week	-0.01	-0.06	-0.23	-0.07	0.00	0.44	-0.04	-0.13										
10	IV30	0.03	0.05	0.22	0.26	-0.01	-0.13	0.12	0.13	-0.22									
11	IV60-IV360	0.01	0.03	0.12	0.31	-0.01	-0.07	0.07	0.10	-0.14	0.51								
12	IV60-StdRet	-0.01	0.03	0.12	-0.19	-0.01	-0.09	0.00	0.10	-0.13	0.00	-0.01							
13	Skew60	0.00	-0.03	-0.07	-0.03	0.01	0.14	-0.02	-0.07	0.09	-0.12	-0.06	-0.17						
14	logME	0.00	-0.06	0.02	-0.17	0.00	0.01	-0.01	0.00	0.03	0.01	-0.10	0.02	0.02					
15	RNSkewness	0.01	0.05	0.14	0.15	0.00	-0.12	0.05	0.06	-0.15	0.18	0.16	0.06	-0.55	-0.21				
16	StkVolume	0.01	-0.04	0.04	0.03	0.00	0.01	0.05	0.00	0.00	0.06	0.02	-0.04	-0.03	0.71	-0.06			
17	OptVolume	0.00	0.01	0.01	0.02	0.00	0.00	0.03	0.00	0.00	0.03	0.02	-0.01	0.03	0.15	-0.09	0.19		
18	OpenInt(p=1)	0.00	0.01	0.01	-0.02	0.01	0.00	0.00	0.00	0.00	0.02	0.01	-0.02	0.04	0.18	-0.11	0.20	0.67	
19	OptBidAsk	0.01	0.02	0.07	0.12	0.00	-0.01	0.10	0.05	-0.04	0.17	0.12	-0.04	0.09	-0.16	-0.01	-0.23	-0.06	-0.08

Table A6 An instrumental variable approach to identifying the channel for returns/order flow predictability. The table shows that order imbalance can predict future option returns through its ability to predict future imbalance rather than volatility. Columns 2 through 5 report the first stage of the 2-SLS regression. The last four columns report different versions of the second stage of 2-SLS. I use six instruments: lag of market-wide order imbalance ("MWOrdImb"), two lags of individual imbalance (OrdImb), lag of order imbalance for short-term options (OrdImb(p=1)) as well as two volatility measures: absolute stock return (AbsStkRet) and adjusted absolute return (RelAbsStkRet). Option returns (Ret) are computed for a delta-neutral straddle portfolio (long) based on the call-put pair which is closest to at-the-money for options with approximately 30 days to expiration. The order imbalance is based on the difference between the number of buy and sell trades, normalized by the total number of trades on a given day (OrdImb) or by the average number of trades in the previous 30 days (AdjOrdImb). RelAbsStkRet is absolute stock returns normalized by its average over previous 50 days. All regressions include a battery of control variables. The absolute t-statistics reported in parentheses are based on robust standard errors clustered by date.

	1^{st} Stage				2^{d} Stage			
	AdjOrd Imb_t	MWOrd Imb_t	AbsStk Ret_t	RelAbs $StkRet_t$	$OptRet_t$			
$MWOrdImb_{t-1}$	0.642	0.632	0.035	1.919				
	(8.03)	(24.10)	(4.33)	(6.06)				
$OrdImb_{t-1}$	0.180	-0.001	0.000	0.014				
	(30.86)	(3.10)	(1.70)	(2.50)				
$OrdImb_{t-2}$	0.088	0.001	0.000	0.022				
	(38.41)	(3.59)	(2.12)	(4.03)				
$OrdImb(p=1)_{t-1}$	-0.010	0.001	0.000	0.020				
	(3.13)	(1.93)	(2.52)	(3.32)				
$AbsStkRet_{t-1}$	0.181	0.006	0.038	-4.174				
	(2.26)	(0.19)	(2.24)	(8.76)				
$RelAbsStkRet_{t-1}$	0.010	0.001	-0.001	0.094				
	(5.56)	(1.94)	(2.83)	(8.27)				
$IV\_AdjOrdImb_t$					0.078	0.043	0.043	0.044
					(12.89)	(15.76)	(15.79)	(12.83)
$IV\_MWOrdImb_t$						0.279	0.279	0.300
						(7.12)	(7.11)	(6.06)
$IV\_AbsStkRet_t$							0.000	
							(0.00)	
IV_RelAbs $StkRet_t$								-0.007
								(0.71)
Other Controls	+	+	+	+	+	+	+	+
R^2	0.04	0.68	0.25	0.10	0.05	0.05	0.05	0.05
N (in 1000s)	1,186	1,215	1,221	1,221	1,175	1,169	1,169	1,169

Table A7 Other significant predictors of option returns. Regressions of future option returns on lagged explanation variables and order imbalances. Column 2 reports standard deviations to facilitate the computation of economic magnitude. Column 4 uses a subsample of 200 stocks with most liquid options (measured as dollar options volume over previous 250 days). Column 5 uses a subsample of options with the bid prices larger than 2 dollars. Column 6 reports regression for two-day-ahead option returns (day *t+1*) as a dependent variable. The last column reports individual regressions with only intercept and the variable itself. Variables are described in Appendix B.”RelOICall” is call options open interest relative to total. ”Weekend” is Friday dummy controlling for weekend returns. “$AbsStkRet_{t-1}$” is absolute stock returns. “$diff(IV)_{t-1}$” is one-day change in implied volatility for short-term options. “OptBidAsk” is the dollar option bid-ask spread. “MeanStkVolume” is median stock volume in the previous 75-days. “IV60-StdRet” is volatility premium measured as different between 60-day ATM implied volatility and historical volatility. “IV(p=1)” is implied volatility for short-term options. “COStdRet - StdRet” is open-close volatility relative to close-to-close volatility. All regressions except the last column control the expiration period and earnings announcement dummies. All variables have subscript “t-1” unless otherwise stated. The absolute t-statistics reported in parentheses are based on robust standard errors clustered by date.

	Std. Dev.	$OptRet_t$	$OptRet_t$ 200Big	$OptRet_t$ Prc. > $2	$OptRet_{t+1}$	$OptRet_t$ Individ.
RelOICall(p=1)	0.18	-0.003	-0.003	-0.002	-0.004	-0.001
		(1.91)	(0.85)	(0.55)	(3.06)	(1.53)
RelVolCall(p=1)	0.30	0.002	0.004	0.003	0.004	0.0
		(3.15)	(1.59)	(2.23)	(2.14)	(0.12)
RelVolPutOTM(p=1)	0.25	0.002	0.004	0.003	0.001	0.001
		(2.54)	(1.45)	(1.75)	(0.60)	(2.09)
Weekend	0.40	-0.013	-0.011	-0.001	0.000	-0.013
		(4.47)	(3.56)	(0.11)	(0.06)	(4.31)
AbsStkRet	0.02	0.162	0.118	0.101	0.070	0.129
		(6.00)	(3.33)	(2.57)	(2.23)	(5.57)
$OptRet_{t-1}$	0.10	-0.042	-0.032	-0.032	-0.004	-0.008
		(7.64)	(3.54)	(3.36)	(0.54)	(1.42)
diff(IV)	0.04	-0.096	-0.043	-0.058	-0.013	-0.072
		(7.09)	(2.10)	(2.91)	(0.78)	(4.88)
OptBidAsk	0.25	0.003	0.011	0.003	0.009	0.004
		(1.49)	(2.90)	(1.31)	(4.20)	(1.70)
$StkRet_{t-2}$	0.03	0.071	0.094	0.132	0.021	0.025
		(2.50)	(2.74)	(3.19)	(0.95)	(0.91)
StkRet6M	0.39	-0.004	-0.004	-0.002	-0.003	-0.005
		(2.63)	(2.35)	(1.51)	(2.02)	(2.93)
MeanStkVolume	0.28	0.008	0.005	0.007	0.002	0.010
		(3.56)	(1.62)	(2.02)	(1.12)	(3.89)
IV60-StdRet	0.16	-0.003	-0.004	-0.010	-0.002	-0.005
		(0.93)	(1.09)	(1.89)	(0.55)	(1.18)
IV(p=1)	0.23	-0.017	-0.007	-0.003	-0.003	0.003
		(3.44)	(1.17)	(0.38)	(0.69)	(0.97)
COStdRet - StdRet	0.03	-0.028	-0.089	-0.083	-0.019	-0.010
		(3.87)	(2.33)	(2.33)	(2.67)	(1.98)
MeanOptVolumeUSD	2.12	0.000	0.001	0.000	0.000	0.0
		(0.64)	(2.11)	(0.20)	(0.56)	(1.07)
Other Controls		+	+	+	+	-
R^2		0.02	0.02	0.03	0.02	
N (in 1000s)		1,253	251	214	1,191	

Table A8 Robustness tests. Regressions of future option returns on lagged explanation variables and order imbalances. Column 2 reports a baseline case with short-term option returns as a dependent variable. Columns 3 and 4 study option returns on day *t+1* (two-days ahead) and day *t+4* (one week ahead). Column 5 uses delta-neutral call returns instead of straddle returns. Column 6 uses a subsample of 200 stocks with most liquid options (measured as dollar options volume over previous 250 days). The last column studies changes during the 2008 crisis. "n_crises" is a dummy which equals to one between August 2007 and January 2009. Option returns are computed for a delta-neutral straddle portfolio (long) based on the call-put pair which is closest to at-the-money. The order imbalance ("OrdImb") is based on the difference between the number of buy and sell trades normalized by the total number of trades on a given day. MWOrdImb is a market-wide order imbalance. All regressions include a battery of control variables. The absolute t-statistics reported in parentheses are based on robust standard errors clustered by date.

	$OptRet_t$	$OptRet_{t+1}$	$OptRet_{t+4}$	$OptRetCall_t$	$OptRet_t$ 200Big	$OptRet_t$
$OrdImb_{t-1}$	0.007	0.004	0.001	0.010	0.009	0.007
	(15.79)	(7.05)	(2.74)	(13.43)	(8.46)	(13.45)
$MWOrdImb_{t-1}$	0.195	0.128	0.101	0.249	0.190	0.155
	(9.32)	(5.74)	(3.61)	(8.25)	(8.35)	(6.79)
n_crises						0.012
						(3.01)
n_crises*$OrdImb_{t-1}$						0.002
						(2.33)
n_crises*$MWOrdImb_{t-1}$						0.119
						(2.28)
Other Controls	+	+	+	+	+	+
R^2	0.02	0.01	0.01	0.02	0.02	0.02
N (in 1000s)	1,132	1,112	1,111	1,132	236	1,132

Table A9 Robustness tests for the microstructure method. The table adds several new variables compared to the previous table and also considers ten-second evaluation period. The new control variables include the contemporaneous and lagged stock returns (adjusted for direction and option delta, $dS_t * |\Delta|/C_t$, $dS_{t-1} * |\Delta|/C_t$), dummies for the first and last hours of trading, earnings announcement day dummy. Due to new control variables (earnings days), stock-day fixed effects have to be replaced with stock fixed effects. The absolute t-statistics reported in parentheses are based on robust standard errors clustered by date. Sample size is 7,684,040 observations for all regressions.

	Five-second price impact		Ten-second price impact	
	Information	Inventory	Information	Inventory
Absolute Delta, $\|\Delta\|$	-0.569	-1.525	-0.855	-1.708
	(18.21)	(39.12)	(19.20)	(38.91)
$\|\Delta\|, if\ \|\Delta\| < 0.4$	-0.119	-0.363	-0.128	-0.388
	(5.88)	(12.46)	(4.71)	(11.35)
$\|\Delta\|, if\ 0.4 \leq \|\Delta\| < 0.6$	-0.121	-0.253	-0.142	-0.273
	(14.51)	(23.06)	(13.27)	(21.24)
$\sqrt{\text{Days to Expiration}}$	-0.017	-0.040	-0.028	-0.045
	(16.93)	(38.64)	(19.22)	(40.06)
Call/Put Dummy	-0.016	0.061	-0.006	0.074
	(4.97)	(13.58)	(1.22)	(13.47)
Option Price, C_t, $	0.005	0.032	0.018	0.040
	(3.09)	(14.36)	(7.21)	(15.96)
Buy/Sell Dummy	0.035	0.160	0.049	0.171
	(5.72)	(26.60)	(7.16)	(25.18)
$\sqrt{\text{Trade Size}}$	0.030	0.017	0.029	0.019
	(38.09)	(18.35)	(27.26)	(17.20)
Trade Size, Contracts	-0.000	-0.000	-0.000	-0.000
	(16.39)	(9.66)	(14.20)	(10.36)
$dS_t * \|\Delta\|/C_t$	0.572	0.082	0.865	0.092
	(44.22)	(8.78)	(57.05)	(9.33)
$dS_{t-1} * \|\Delta\|/C_t$	0.345	0.229	0.407	0.236
	(33.11)	(22.71)	(33.66)	(22.89)
Last Hour Dummy	0.008	0.019	0.020	0.027
	(2.04)	(3.48)	(3.11)	(3.81)
First Hour Dummy	0.019	0.039	0.035	0.048
	(5.64)	(8.71)	(7.41)	(8.67)
Earnings Day Dummy	0.045	0.034	0.078	0.036
	(2.50)	(2.25)	(2.99)	(2.16)
R^2	0.04	0.01	0.05	0.01