3-2014

Aggregating information in option transactions

Richard Holowczak
Singapore Management University, JIANFENGHU@smu.edu.sg

Jianfeng HU
Singapore Management University, JIANFENGHU@smu.edu.sg

Liuren WU
CUNY Bernard M Baruch College

Consolidating Information in Option Transactions*

Richard Holowczak

Zicklin School of Business, Baruch College

Jianfeng Hu

Zicklin School of Business, Baruch College and the Graduate Center, CUNY

Liuren Wu

Zicklin School of Business, Baruch College

This version: March 15, 2011; First draft: December 6, 2009

Abstract

Underlying each stock trades hundreds of options at different strike prices and maturities. The order flow from these option transactions reveals important information about the underlying stock price movement and its volatility variation. How to aggregate the trade information of different option contracts underlying the same stock presents an interesting and important question for developing microstructure theories and price discovery mechanisms in the derivatives markets. This paper takes options on QQQQ, the Nasdaq 100 tracking stock, as an example and examines different order flow consolidation schemes in terms of their effectiveness in extracting information about the underlying stock price movement and its volatility variation. The analysis shows that an effective consolidation scheme shall account for each contract's different exposure to the stock price and volatility movements. The scheme should also accommodate concerns on liquidity and interference from other potential risk dimensions, such as market crashes and long-term versus short-term volatility factors. Based on our proposed consolidation scheme, we identify significant relations, both contemporaneous and predictive, between the appropriately aggregated options order flows and the stock returns and return volatilities. In particular, the aggregated stock buy pressure positively predicts future stock returns and the aggregated volatility buy pressure positively predicts future changes in return volatilities up to five minutes.

JEL Classification: G14, G12, G13.

*The authors thank Turan Bali, Armen Hovakimian, Robert Schwartz, and seminar participants at Baruch College, 2010 Northeast Business and Economics Association(NBEA) annual conference, FDIC-Cornell-University of Houston Derivative Securities and Risk Management Conference for comments. We welcome comments, including references we have inadvertently overlooked. Jianfeng Hu thanks CUNY Graduate Center for graduate research grant. Richard.holowczak@baruch.cuny.edu (Holowczak), jianfeng.hu@baruch.cuny.edu (Hu) and liuren.wu@baruch.cuny.edu (Wu).

Keywords: Options order flow; information aggregation; delta; vega; lead-lag relations; price discovery; OPRA; QQQQ.

Consolidating Information in Option Transactions

This version: March 15, 2011; First draft: December 6, 2009

Abstract

Underlying each stock trades hundreds of options at different strike prices and maturities. The order flow from these option transactions reveals important information about the underlying stock price movement and its volatility. How to aggregate the trade information of different option contracts underlying the same stock presents an interesting and important question for developing microstructure theories and price discovery mechanisms in the derivatives markets. This paper takes options on QQQQ, the Nasdaq 100 tracking stock, as an example and examines different order flow consolidation schemes in terms of their effectiveness in extracting information about the underlying stock price and volatility movements. The analysis shows that an effective consolidation scheme shall account for both the liquidity of each option contract and the contract's particular exposure to stock price and volatility movements. Based on our proposed consolidation scheme, we identify significant relations between the appropriately aggregated options order flows and the ex post stock returns and return volatilities. In particular, the aggregated stock buy pressure positively predicts future stock returns and the aggregated volatility buy pressure positively predicts future changes in return volatilities up to five minutes.

JEL Classification: G14, G12, G13.

Keywords: Options order flow; information aggregation; delta; vega; lead-lag relations; price discovery; OPRA; QQQQ.

Contents

1. Introduction

In the absence of market frictions and under the geometric Brownian motion stock price dynamics assumed in Black and Scholes (1973) and Merton (1973), options can be perfectly replicated by a portfolio of a risk free bond and the underlying stock. Option trading is thus redundant. In reality, however, the market shows a strong demand for options for two major reasons. First, the risks in the stock market cannot be completely spanned by the stock trading alone. For example, the presence of discontinuous stock price movements of random sizes necessitates the inclusion of options across a whole spectrum of strikes to span the jump risk (Carr and Wu (2004)). The presence of stochastic volatility (Engle (2004)), on the other market, makes the options market the *de facto* market for trading volatility risks (Carr and Wu (2009)).

The second major reason for options trading is informational. Even in the absence of stock price jumps and stochastic volatility, investors may choose to trade options to gain exposure to the stock given the high leverage provided by options (Black (1975)). Furthermore, informed traders may prefer the options market because they can better hide themselves among the multiple option contracts available on one security (Easley, O'Hara, and Srinivas (1998)). Trading options also allows investors to express their view in volatility, which cannot be fully expressed with stock trading alone. On the other hand, the transaction cost on options is usually much higher than on the underlying stock. Thus, only when the perceived information advantage is large enough do the benefits of high leverage and multiple contract availability overshadow the large transaction costs (Holowczak, Simaan, and Wu (2006)).

A long list of studies have investigated the information flow between the options market and stock market.[1] The challenge remains on how to effectively aggregate the information in the multiple option contracts underlying the same stock. When the underlying stock price moves, no arbitrage dictates that the prices on all the option contracts underlying this stock will move accordingly. When an options market maker takes on a position in any of the options underlying the same stock, the market maker will use the same stock to perform delta hedging to remove the underlying stock exposure. Hence, it is important to aggregate the information from the diverse option transactions at different strikes and maturities before one links the options transactions to stock price movements.

Most existing studies either use only one pair of option contracts (e.g., Chan, Chung, and Fong (2002)

[1] See, for example, Manaster and Rendleman Jr. (1982), Bhattacharya (1987), Anthony (1988), Stephan and Whaley (1990), Finucane (1991), Chan, Chung, and Johnson (1993), Easley, O'Hara, and Srinivas (1998), Jarnecic (1999), Chan, Chung, and Fong (2002), Chakravarty, Gulen, and Mayhew (2004), and Holowczak, Simaan, and Wu (2006).

and Holowczak, Simaan, and Wu (2006)) or regard different contracts as equally informative (e.g., Easley, O'Hara, and Srinivas (1998), Chakravarty, Gulen, and Mayhew (2004), Cao, Chen, and Griffin (2005), and Pan and Poteshman (2006)) to simplify the problem.[2] Picking one pair of contracts while discarding all the others amounts to throwing away a large amount of information, and can potentially distort the estimated relations. One can think of the case where the chosen option contract has a small transaction while most other options experience large transactions pointing to the opposite direction for the stock price movement. In this case, the large transactions of the omitted option contracts, rather than the small transaction of the chosen contract, are likely to dictate the direction of the stock price movement. Equal weighting can be equally problematic as informed traders do not randomly pick an option contract to trade. Instead, they will consider market depth, liquidity, and leverage to optimize their contract allocation.

Another important but rarely raised question is how to aggregate the order flow across trades with different sizes. While most researchers focus on option volumes (Easley, O'Hara, and Srinivas (1998), Chan, Chung, and Fong (2002), Chakravarty, Gulen, and Mayhew (2004), Cao, Chen, and Griffin (2005), Pan and Poteshman (2006), Ni, Pan, and Poteshman (2008)), some scholars just count number of trades to construct "option buy pressures" (Bollen and Whaley (2004) and Holowczak, Simaan, and Wu (2006)). By counting number of trades, they presume that trades are equally informative across different sizes. This method is likely to overstate the impact of very small trades because small odd-lot trades are generally regarded as non-informative retail trades by options market makers. On the other hand, aggregating trading volumes assumes that the information content is linear to trade size, and this assumption can overstate the informativeness of large trades. First, very large trades are often pre-negotiated in the upstairs market and thus may not contain the most updated information. Second, informed traders often split their orders to disguise themselves among uninformed orders. For example, Anand and Chakravarty (2003) find that medium size option trades are often used to achieve "stealth trading" when trading volume is high.

In this paper, we propose a mechanism to aggregate option transactions across all strikes and maturities on the same stock, and we test its effectiveness against four alternatives in terms of their effectiveness in extracting information about future stock price and volatility movements. In extracting the information on stock price movement, the first consideration is the stock price exposure. A call option has positive stock price exposure and a put option has negative stock price exposure. Accordingly, aggregations of buy and

[2]More recently, Bollen and Whaley (2004) examine the impact of absolute delta weighted option order flows on implied volatility functions. Ni, Pan, and Poteshman (2008) use price-scaled vega weighted option volumes to predict realized volatilities in the cross section.

sell orders on call and put options should take on opposite signs. A standard measure for the stock price exposure is the delta of the option, which measures how much the option price moves when the underlying stock price moves by one dollar. The second consideration is leverage. Given limited capital and private information, an investor would want to maximize its delta exposure per dollar spent on the contract. The delta of an option scaled by the option's value represents the stock exposure per dollar spent. Finally, the investor must account for the different transaction costs on different options contracts in terms of both bid-ask spreads and market impacts. The options market liquidity concentrates on short-term near-the-money options. Although the stock exposure per dollar spent is the highest for far out-of-the-money options, the high bid-ask spread relative to the option value makes these contracts prohibitively expensive to trade. We combine all three considerations to generate an aggregate stock buy pressure (ASBP) for option transactions. We show that this ASBP measure generates significant predictions on future stock price movements.

We also propose an aggregate order flow measure that reveals the information in the underlying volatility. In this case, we focus on the volatility risk exposure, or vega, of each option contract instead of its delta exposure. We combine the vega exposure with the leverage and liquidity considerations to construct an aggregate volatility buy pressure (AVBP) for option transactions. We find that this AVBP measure predicts future stock volatility as measured by the standard deviation of realized stock returns.

To address the second question on how to aggregate across trade size, we directly test the effectiveness of three candidate measures: number of trades, volume, and the natural logarithm of trading volume. We find that the log volume outperforms the rest regardless of how we aggregate order flows across strikes and maturities. Under the natural log transformation, odd-lot transactions are heavily discounted for the information content. The information content is assumed to increase with the trade size, but flattens out up to a certain size.

Many studies investigate the information flow between the options market and the stock market, often with conflicting findings. Early studies such as Manaster and Rendleman Jr. (1982) and Bhattacharya (1987) find that the options market reveals information about the underlying security prices. Easley, O'Hara, and Srinivas (1998) do not find the option prices to be informative but they find that signed option volumes are informative about future stock prices although the directions are not as expected. Using the information share approach developed by Hasbrouck (1995), Chan, Chung, and Fong (2002) and Chakravarty, Gulen, and Mayhew (2004) find that the stock market leads the option market in price discovery. Holowczak, Simaan, and Wu (2006) find that the statistical significance of price discovery varies with option trading intensity and sidedness. Using a unique data set, Pan and Poteshman (2006) find option call-put volume ratios predict future stock returns; and Ni, Pan, and Poteshman (2008) find that daily dollar-vega weighted order flow predicts future realized volatility.

Our work contributes to the literature by providing a systematic analysis on the aggregation of option transactions across different strikes, maturities, and size. One cannot possibly obtain robust results on the information flow between the options market and the stock market without first resolving the aggregation issue. We focus our analysis on the aggregation of option transactions, but the same mechanism can also be applied to aggregations of option quotes. Our finding of significant relations between option order flows and stock market movement in high frequency public data also provides direct empirical evidence for information trading on options market.

The rest of the paper is organized as follows. Section 2 discusses various order flow aggregation approaches and propose our own aggregation method based on balanced considerations of exposure, leverage, and transaction cost. Section 3 describes the data we use to test the effectiveness of different aggregation schemes. Section 4 reports the findings from the comparative analysis. Section 5 further explores the power of aggregate option buy pressures under changing market conditions. And section 6 concludes.

2. Aggregating Options Order Flow

2.1. Determining the direction of each transaction

The transaction data base normally does not have a flag on whether the non-market maker party in a trade is the buyer or the seller. In order to determine the option order flow, we follow Lee and Ready (1991) to classify trades into three categories: buyer-initiated, seller-initiated, and unclassified. The signing algorithm is as follows. If a trade price is above the last effective mid quote, it is classified as buyer-initiated. If a trade price is below the mid quote, it is classified as seller-initiated. If a trade price falls exactly on the mid quote and is higher than the last different trade price, it is classified as buyer-initiated. If a trade price falls exactly on the mid quote and is lower than the last different trade price, it is classified as seller-initiated. Everything else is unclassified.

With this signing algorithm, we are able to classify most trades, leaving only 0.86% of the total trades in the unclassified category. These trades normally occur in market opens when there are no valid quotes or last different prices. We discard these unclassified trades from our analysis.

Unlike Lee and Ready (1991), however, we examine the last (t) effective quotes on the same exchange rather than the five-second proceeding ($t-5$) quotes. The reporting lag is not needed for the options data because we find that the proportion of trade-through trades (with the price outside the quote bounds) increases with the time lag and using zero delays generates the largest proportion of trades exactly on the bid or the ask.

2.2. Aggregating the net buy pressure of each contract

For each option contract, we define the net buy pressure as the buy transactions minus the sell transactions over a certain time horizon, and we use $CBP(K,T)$ to denote the net buy pressure from a call option at strike K and expiry T, and use $PBP(K,T)$ to denote the buy pressure of a put option at strike K and expiry T.

Aggregating horizon is not a trivial choice in microstructure studies. As market efficiency improves over time, the speed of information dissemination across markets also increases. A long aggregating horizon may not be sensitive enough to capture the information content in option transactions as trade imbalance gradually dies out. A short aggregating horizon, on the other hand, can increase the underlying price sensitivity to option buy pressures. However, it may also inflate the effect of noise in trading, especially when trading is inactive. There is clearly a tradeoff between a long observation horizon and a short one. In this paper, we choose an ETF as the underlying security and one may argue that informed trading is not likely to occur for an ETF and its options because of less information asymmetry at portfolio level. While it is true that the macro variables driving the ETF's price and volatility are more transparent than single stock fundamentals, there can be as much, if not more, liquidity information in the market, which dictates the movement of the underlying price and volatility in short-term intervals ranging from a few seconds to a few minutes. Therefore, we choose a relatively short aggregating horizon of one minute to conduct our analysis. The observation length is less than the traditional length of five minutes in, for example, Stephan and Whaley (1990) and Easley, O'Hara, and Srinivas (1998). The selection is not likely to overstate the noise impact from inactive trading periods because in our sample there are 12.07 option trades in an average one-minute observation period, more than 2.57 option trades in an average five-minute observation period in the sample of Easley, O'Hara, and Srinivas (1998). For robustness check, we also repeat our analysis for aggregating horizons at five seconds, ten seconds, thirty seconds, five minutes, and fifteen minutes. Our main results remain unchanged qualitatively.

When there are multiple buy and sell transactions of different size within the same time period, we must decide on how the transactions of different size are aggregated. We consider three alternative ways of aggregation on each option contract.

1. **Number of trades**, where each transaction is treated as one unit when being aggregated, regardless of the size of the transaction. Several empirical studies on stock market microstructure have found that number of trades is more informative than trade volume, e.g.,Jones, Kaul, and Lipson (1994), Ané and Geman (2000), and Izzeldin (2007).

   However, using number of trades is likely to overstate the importance of very small trades in options. In the stock options market, orders of five contracts or less are referred to as “odd lots,” and are generally considered as non-informative retail trades. Indeed, at some options exchanges such as the International Options Exchange, a large proportion of the odd-lot trades are rewarded to the primary market maker as a compensation for the extra responsibilities (Simaan and Wu (2007)).

2. **Trade volume**, which amounts to assume that the importance of each trade is proportional to its size. In the stock market, very large trades are often negotiated in the upstairs market and are put into the print at a later time. As a result, the reported large trade tend to be lagged report and is thus not as informative. The same practice also happens on the options exchanges. Thus, using trading volume is likely to overestimate the information value of very large trades. Furthermore, with the explosion of electronic trading, large institutional trades are increasingly being sliced into smaller orders in the stock market as a way to mitigate the market impact (Anand and Chakravarty (2003)). Such strategic order placing technique is also applied in the options market now, quite often referred to as ”board surfing” by practitioners. Informed traders are more likely to slice orders to hide their trading intention. The total trade volume is not able to capture the order slicing effect and may understate the information content in the sum of small to medium sized trades.

3. **Log volume**, where the contribution of each trade is proportional to the natural logarithm of the trade size. By taking natural logs, we apply a discount to the odd-lot trades, and allow the information value to increase with increasing trade size. Nevertheless, the marginal contribution declines as the trade size further increases. Furthermore, when an institution deliberately split a large transaction into several consecutive medium sized transactions, the aggregated information value becomes higher than

the value from a single, non-split order.[3]

2.3. Aggregating net buy pressure across different option contracts

To aggregate the net buy pressure across different option contracts, we must consider the following factors:

1. **Exposure**: If we intend to infer the underlying stock price movement from the options order flow, a buy pressure from a call option is likely to have the opposite effect to a buy pressure from a put option. Indeed, according to put-call parity, buying a call option while selling a put option at the same maturity and strike price is equivalent to long a forward contract on the underlying stock. It is this idea that has motivated most empirical research to find call and put option pairs at the same maturity and strike, and compute the net stock buy pressure from the pair as the call buy pressure minus the put buy pressure,

$$SBP(K,T) = CBP(K,T) - PBP(K,T), \tag{1}$$

   where $SBP(K,T)$ denotes the net stock buy pressure originated from the options transactions at strike K and maturity T. While the put-call parity is only applicable to European options, the idea can nevertheless be borrowed to aggregate order flows from American-style stock options.

   On the other hand, if the objective is to infer the underlying stock volatility movement, a European call option and a European put option at the same maturity and strike share the same volatility exposure as they have the same time value. Thus, we propose to aggregate the volatility buy pressure at each strike and maturity as,

$$VBP(K,T) = CBP(K,T) + PBP(K,T), \tag{2}$$

   where $VBP(K,T)$ denotes the net volatility buy pressure originated from the options transactions at strike K and maturity T. Again, potentially early exercise on American options can induce an asymmetric effect on the call and the put options at the same strike and maturity. We ignore this usually small effect on the order flow aggregation.

2. **Liquidity**: Options orders are mostly concentrated at short maturities and at strikes close to the spot level. These options contracts also tend to have narrower bid-ask spreads. Thus, everything else

[3] A potential problem with log(volume) is that it regards one-lot trades as completely uninformative. Alternatively. we have also used log(volume+1) to incorporate the impact of one-lot trades and find that log(volume+1) does not perform better than log(volume) in aggregating option buy pressures.

equal, informed traders are likely to allocate more of the capital to the most actively traded options to mitigate market impact. On the other hand, transactions at illiquid strike/maturity regions such as deep out of the money and/or at very long maturities can be motivated by other considerations instead of information on the stock or volatility.

3. **Leverage**: Given limited capital, an informed investor may want to maximize its exposure (stock price or volatility) per dollar spent on the contract. Far out-of-the-money options have low stock price and volatility exposures per contract, but their stock price and volatility exposures per dollar spent are actually very high because of the low dollar value of the contract. However, margin requirements can mitigate the effect of leverage, especially for writing options. For example, when a customer sells a far out-of-the-money put option, the customer is required to hold a margin greater than the sales receipt.

4. **Other considerations**: Investors can buy or sell options for other considerations. For example, institutional investors often buy far out-of-the-money put options to hedge against market crash risk. Investors can also use calendar spreads to trade on the variation of the volatility term structure, allowing for potentially separate movements for short-term and long-term volatilities. Thus, when one uses the Black-Merton-Scholes vega to denote the volatility exposure, the vega at different strike and maturity regions can be driven by different stochastic factors.

In developing aggregation schemes for the stock and volatility buy pressures, we strive to achieve a balance among the different considerations. Formally, we construct the aggregate stock buy pressure (ASBP) as,

$$ASBP = \sum_{j=1}^{N} \frac{1}{M_j} SBP(K_j, T_j) n(d_j), \tag{3}$$

where N denotes the number of options contracts, $M_j = \max(1, 12 \times T_j)$ denotes a truncated maturity measure (in months) for the option contract, and $n(d_j)$ denotes the probability density function of a standard normal variable, with

$$d_j = \frac{\ln(F_j/K_j) + \frac{1}{2}\sigma^2 T_j}{\sigma\sqrt{T_j}}, \tag{4}$$

and σ being some return volatility estimate for the underlying stock. In application, we use an average implied volatility estimate from the previous day to proxy σ.

The aggregation in (3) combines all four considerations. First, the stock buy pressure at each strike and maturity (K_j, T_j), defined in (1), recognizes the opposite stock price exposure for the call and the put options at the same strike and maturity, and cancels out the contribution of volatility exposures. Second, across different strikes, the $n(d_j)$ weighting in (3) puts more weight on near-the-money options because near-the-money options tend to be more liquid than out-of-the-money options. Third, across different maturities, the option value scales approximately in the order of $\sqrt{T}$. Through the $1/M_j$ weighting, we divide the buy pressure by maturity to discount the contribution of longer-term contracts for liquidity concerns. We convert the maturity scaling in months and set the minimum to one month to avoid extreme weighting for options at very short maturities.

The aggregate volatility buy pressure (AVBP) is constructed analogously,

$$AVBP = \sum_{j=1}^{N} \frac{1}{\sqrt{M_j}} VBP(K_j, T_j) n(d_j). \tag{5}$$

First, at each strike and maturity, we sum the call and the put buy pressure to generate a volatility buy pressure in (2) that is relatively insensitive to directional movements in the stock price. Second, along the strike dimension, we use the $n(d_j)$ weighting to assign more weights to near the money order flows. Third, along the maturity-dimension, we use $1/\sqrt{M_j}$ maturity scaling to balance out the increasing volatility exposure with the decreasing liquidity as maturity increases. By focusing more weights on short-term near-the-money options, we also reduce the impact of other factors such as crash risk and long-term volatility risk to contaminate our measure for short-term market return volatility.

2.4. Alternative aggregating schemes

We compare the information contents of our proposed aggregate stock and volatility buy pressures to three alternatives that have been used in the literature.

1. **One strike-maturity point.** This approach has been applied in in Chan, Chung, and Fong (2002) and Holowczak, Simaan, and Wu (2006). In our implementation, we pick the stock and volatility buy pressures at the strike and maturity point with the highest number of option trades. The chosen strike-maturity point is always near the money and ay short maturity. In total, this one point accounts for 24% of the total number of trades in our sample.

We have also experimented with the most number of calls and the most number of puts as the selection criterion. The results are similar.

2. **Equal weighting:** Equal weighting has been used in, for example, Easley, O'Hara, and Srinivas (1998),Chakravarty, Gulen, and Mayhew (2004), Cao, Chen, and Griffin (2005), and Pan and Poteshman (2006). The aggregation is the sum of option buy pressures of all individual contracts.

3. **Greek weighting**: The aggregate stock buy pressure is computed as the delta weighted sum of buy pressures on each option. The aggregate volatility buy pressure is computed as the vega weighted sum of the buy pressure on each option. The delta and vega are computed using the Black-Merton-Scholes formula with an average implied volatility as the volatility input. This weighting scheme focuses squarely on the risk exposure of each option contract based on the Black-Merton-Scholes model and ignores other considerations such as liquidity, leverage, and other risk dimensions. If liquidity is not a concern and the underlying stock moves according to the Black-Merton-Scholes model, this greek weighting should be the right choice.

4. **Greek per dollar weighting:** The greek weights (delta for stock buy pressure and vega for volatility buy pressure) are divided by the dollar value of the option contract to adjust for the leverage consideration. Ni, Pan, and Poteshman (2008) use this weighting scheme to predict the future realized volatility. This weighting method also assigns more weight to out-of-the-money options than to in-the-money options. Chakravarty, Gulen, and Mayhew (2004) find that out-of-the-money options have higher information share in price discovery analysis than in-the-money options.

2.5. Performance measure

In order to evaluate the effectiveness of our aggregate option buy pressures, we look at both contemporaneous and forecasting correlations of aggregate option buy pressures and the underlying price and volatility movements. The squared univariate correlation is equivalent to the R-squared of an Ordinary-Least-Squared regression. Thus we can use the magnitude of the univariate correlations to indicate the effectiveness of different aggregation option buy pressures. If option trading does not contain information about the underlying market, we expect the aggregate option buy pressures to be insignificantly correlated with stock returns and volatilities. For the purpose of detecting information based trading in the options market, one might think that the forecasting power of aggregate option order flows should be given more attention than the contemporaneous relations. However, the contemporaneous correlations reflect the impact of option trading on the stock market and are also important.

Alternatively, we also run the following multiple time series regressions to examine the impact of aggregate option buy pressures over time:

$$Ret_t = \alpha 0 + \sum_{i=0}^{k} \alpha 1_k ASBP_{t-k} + \varepsilon_t, \tag{6}$$

and

$$Vol_t = \beta 0 + \sum_{i=0}^{j} \beta 1_k AVBP_{t-j} + \xi_t, \tag{7}$$

where Ret_t denotes the return at time t, Vol_t denotes the realized volatility at time t, and ASBP and AVBP are aggregate stock and volatility buy pressures, respectively. We test the null hypothesize of redundant options trading, i.e. $\alpha 1$ and $\beta 1$ are 0 at all lags. The regression results do not change our conclusion about the effectiveness of different aggregation methods. For reporting simplicity, we only report the time series regression results for the most efficient aggregation method in univariate correlation analysis.

3. Data

3.1. Sample selection

The options data are from the Option Price Reporting Authority (OPRA), which records every option quote and trade message across all option exchanges in the United States. The underlying security price data comes from New York Stock Exchange Trade and Quote (TAQ) database.

We perform our analysis on one underlying security, the NASDAQ 100 index tracking stock (QQQQ). Options on QQQQ are traded on all options exchanges in the United States and are among the most actively traded stock options (Holowczak, Simaan, and Wu (2006)). Our sample covers 231 trading days from February 1st to December 29th, 2006. There are 1,572,865 trade records during this sample period. We filter the trading records by excluding off-hour trades, trades that happen within the first 15 minutes of the market open and the last five minutes before the market close, and trades on options that expire within 10 calendar days. These filtering reduces the sample to 1,087,778 trade records, which average about 4,709 trades per day.

The trade size averages at 65 contracts, with a median of five contracts, a minimum of one contract and a maximum of 275,000 contracts. The transactions are heavily concentrated at small sizes. In particular, about 51% of the transactions are "odd lots" (with five or fewer contracts). Figure 1 plots the histogram of the trade size represented in natural logarithms of the number of lots. Even after taking the logs, the trade size distribution is still heavily skewed to small sizes.

[Figure 1 about here.]

Table 1 further classifies the option transactions in terms of call and put options and across different moneyness and maturity regions. Among the 1,087,778 trades, 506,948 of them, or 46.6%, are call options and 580,830 trades, or 53.4%, are put options. The put option trade sizes average higher than the call option trade sizes. As a result, the average daily trading volume for the put options at 184,494 lots represents a larger percentage (60%) of the total trading volume.

When we classify the option transactions in terms of moneyness, 44.19% of the transactions have strikes close to the spot, with the delta of options between 37.5% and 62.5%. 41.26% of the transactions are out-of-the-money (OTM) options with the delta below 37.5% and only 14.55% are in-the-money (ITM) options with the delta above 62.5%.

Table 1 also shows that the majority (79.91%) of the transactions are short term options expiring in less than two months. The trade direction is slightly imbalanced with 52.06% of trades initiated by buyers in the entire sample. Call options are more balanced with 49.32% trades being buyer-initiated and put options are more imbalanced with 54.46% trades being buyer-initiated. The proportions in table 1 are computed with number of trades.

[Table 1 about here.]

3.2. Main variables

We calculate returns and volatilities of the underlying security QQQQ as our dependent variables. Raw returns are calculated as the difference between the log mid quotes of the National Best Bid and Ask (NBBO) at the beginning and the end of each one minute interval during trading hours. Table 2 reports that the mean of the one-minute raw return is -0.006 basis points (bp). Bid ask bounce of the stock price can cloud the validity of our findings through serial correlations. Following Easley, O'Hara, and Srinivas (1998), we use

MA(1) process to remove the autocorrelation of the returns for each trading day. We name the residual of MA(1) excess return and use that as the dependent variable to examine the impact of ASBP. The first order serial correlation in the return series is reduced from -0.013 to -0.007 after performing the MA(1) procedure. The mean of one minute excess return is -0.002 bp in our sample with standard deviation of 4.516 bp.

To measure stock return volatility, we calculate the standard deviation of second by second returns within each one-minute observation interval and standardize it to annual volatility assuming the stock price follows a log normal process. Table 2 reports that the mean of this annualized volatility is 3.035%. We find using other volatility measures such as (high-low)/average and option implied volatility does not change our results qualitatively.[4]

[Table 2 about here.]

Table 2 also reports the mean and standard deviation of each aggregate option order flow we construct in section 2. We use three month implied volatility from Bloomberg and assume 0 interest rate and dividend rate to compute the B-S-M Greeks. The mean aggregate stock buy pressures are generally negative and the mean aggregate volatility buy pressures are all positive although none is statistically different from 0. The consistency in order flow directions shows the aggregate option buy pressures are not illy measured. We apply Augmented Dickey and Fuller (1979) test for all buy pressure variables with lags of 20. Not surprisingly, the p-values are all less than 0.001, strongly rejecting the null hypothesis of non-stationarity.

4. Comparative analysis

4.1. Aggregate stock buy pressures

An informed trader can trade options to profit from directional movement of the underlying security price. If our aggregate stock buy pressures are correctly constructed, they are supposed to be positively correlated with the excess returns. Easley, O'Hara, and Srinivas (1998) find option volumes are informative about stock prices. However, they find a significantly negative coefficient of the contemporaneous five-minute option volume instead of the hypothesized positive one. We report the correlations of excess returns and ASBP in table 3. Contradicting their result, we find that all contemporaneous correlations are positive and significant (at 1% level) as expected in our sample as shown in panel A of table3, supporting the existence

[4]Results for alternative volatility measures are available upon request.

of information based trading in options market.

[Table 3 about here.]

The second result emerging from Panel A of table 3 is that using log(volume) always outperforms the other aggregations across trade size regardless of what method is used to aggregate across strikes and maturities. Between number of trades and volume, number of trades has twice to three times the coefficients of those of volume except for *one_pair* and *delta_per_dollar*, which generate close coefficients for both aggregation methods.

Among the aggregation methods across strikes and maturities, *delta_per_share* outperforms the rest with *equal_weighting* and our proposed *HHW* closely tied in the second place. The differences between these three are small. *delta_per_dollar* comes after with a larger gap and *one_pair* always has the smallest coefficient. A surprising result is the underperformance of *delta_per_dollar* to *delta_per_share* because the former is expected to capture the leverage effect which should be one of the main reasons for information trading in options market. One possible explanation is that *delta_per_dollar* aggregation overstates the effect of penny trades of deeply OTM options and ignores the liquidity effect which is more important. The largest coefficient (0.411) belongs to ASBP from *delta_per_share* and log(volume). And the smallest coefficient (0.050) belongs to ASBP from *one_pair* and number of trades.

To show the results are not driven by our choice of the observation horizon at one minute, we perform the same analysis for different observation horizons ranging from 5 seconds to 15 minutes. The contemporaneous correlations are plotted in figure 2. Generally we see that the contemporaneous correlations increase in the observation length and the main results from Panel A of table 3 holds in different observation lengths, i.e. *delta_per_share* and log(volume) outperform the other aggregation methods in their categories. Figure 2 also shows that different aggregation methods across strikes and maturities converge to a great extent under log(volume) aggregation except for *one_pair*. It seems the contemporaneous correlation is more sensitive to the aggregation method across trade size than the aggregation method across strikes and maturities.

[Figure 2 about here.]

Now let us turn to forecasting the excess returns with aggregate stock buy pressures. Panel B of table 3 reports the correlations of the excess returns and one period lagged (t-1) ASBP. All significant correlations are positive, consistent with our hypothesize. Comparing the aggregation methods across sizes first, we can see that number of trades is not informative about future returns as the correlations are all insignificant. Between the other two methods, volume weighting seems to lead the race. Four out of five measures (except *delta_per_dollar*) aggregated from option volume generate significant and positive correlations to future excess returns although the correlations are much smaller than corresponding contemporaneous correlations. Log(volume) weighting generates two significant correlations for ASBP of *one_pair*(statistically significant at 1%) and *HHW*(statistically significant at 10%).

Among the aggregation methods across strikes and maturities, *HHW*and *one_pair*have two significant correlations of ASBP aggregated from volume and log(volume). *delta_per_share*and *equal_weighting*have one significant correlation of ASBP aggregated from volume only. And *delta_per_dollar*ASBP again does not have any significant correlation with the excess returns. Volume weighted ASBP aggregated with both *HHW*and *one_pair*have the largest correlation of 0.011 with succeeding one-minute excess returns.

Figure 3 plots the correlations in different observation horizons. At short-term observation horizons (less than five minutes), we find the correlation pattern is consistent with the results documented above. However, at longer horizons, the aggregate stock buy pressures give very different predictions and some are even negatively correlated with future excess returns. Although not reported, the correlations generally become statistically insignificant after the observation length increases to five minutes.

[Figure 3 about here.]

What can we say about the effectiveness of these aggregate stock buy pressures then? First of all, among aggregation methods across strikes and maturities, our proposed *HHW*ASBP persistently generates large and significant correlations regardless of what method is used to aggregates across sizes. *delta_per_share*and *equal_weighting*ASBP also have comparable contemporaneous correlations. However, they do not have the same predictability as *HHW*ASBP. *one_pair*works well in predicting future excess returns. However, this extreme filter discards too much information compared to other aggregations and does not reveal the contemporaneous impact on the stock market as well. Considering both contemporaneous and forecasting relations, we find our *HHW*method should be preferred to the other aggregation methods. Second, among aggregation methods across sizes, log(volume) works best in contemporaneous correlations and volume weighting excels in forecasting. Which aggregation to choose then depends on the nature of particular research questions.

4.2. Aggregate volatility buy pressures

Volatility information is also profitable in the options market and we expect the aggregate volatility buy pressures constructed in section 2 to be positively correlated to the volatilities of underlying returns. Similar to the test of ASBP, we report both contemporaneous and forecasting correlations of AVBP with volatilities in table 4.

[Table 4 about here.]

Panel A of table 4 reports the contemporaneous correlations between AVBP and underlying volatilities. Comparing aggregation methods across strikes and maturities, *HHW* AVBP always have positive and significant correlations regardless of what method is used to aggregate across sizes and the correlations are generally larger than the other correlations in the same row. *one_pair*, *equal_weighting*, and *vega_per_share* AVBP each has two out of three correlations being positive and significant. *vega_per_dollar* AVBP, however, has only one negative and significant correlation with contemporaneous volatilities. Comparing correlations across rows, log(volume) and volume weighted AVBP each has four out of five correlations being significant while only two out of five AVBP constructed from number of trades are positive. Log(volume) weighted AVBP also has larger correlations than the other two methods except in the third column where *vega_per_share* is used. The largest correlation (0.018) in panel A belongs to *HHW* and log(volume) weighted AVBP. Figure 4 plots the correlations in different observation horizons and we can see that the result is not sensitive to the choice of observation horizon. The results here lead us to believe that *HHW* and log(volume) weighted AVBP is the right way to extract the contemporaneous volatility information in the options market.

[Figure 4 about here.]

Next we examine the predictability of AVBP in panel B of table 4. Across rows, number of trades does not generate informative AVBP about future volatilities. Both volume and log(volume) weighted AVBP have some positive and significant correlations with future volatilities but log(volume) weighted AVBP always has larger and more significant correlations. Across columns, *HHW* AVBP has the largest and most significant correlations. *one_pair* and *equal_weighting* AVBP generate positive and significant correlations but the magnitude is smaller. *vega_per_share* AVBP does not generate significant correlation for volume weighting and *vega_per_dollar* AVBP is not informative. Figure 5 plots the forecasting correlations of AVBP in different observation horizons. It is clear that the result is not sensitive to the choice of observation horizon except that *one_pair* volume weighted AVBP loses forecasting power once the observation horizons reaches 5 minutes. The evidence presented here shows that *HHW* and log(volume) weighted AVBP has the greatest forecasting power of future volatilities. The patterns of contemporaneous and forecasting relations of AVBP and volatilities are consistent. Therefore, we are confident to say that when analyzing the impact of options trading on the volatility of stock returns, *HHW* and log(volume) weighted AVBP should be preferred to other aggregations.

[Figure 5 about here.]

As a short summary of the results produced so far, we examine the correlations between aggregate option order flows and the underlying returns and volatilities and find that our proposed *HHW* ASBP and *HHW* AVBP are informative about both contemporaneous and future market movement. Log(volume) weighting generally outperforms volume and number of trades in constructing informative option buy pressures except that volume weighted ASBP better predicts future returns than log(volume) weighted ASVP. The results are robust to alternative observation horizons.

4.3. Robustness

The significant correlations reported in previous subsections are often small and one may worry that the impact of abnormal data points can be huge. In this subsection, we repeat the analysis above using winsorized variables to remove the impact of outliers. The treatment should improve the performance of number of trades and volume weighted buy pressures more than that of log(volume) weighted buy pressures because the logarithm mechanically smoothes the data already. We winsorize all dependent and independent variables at top and bottom 0.1% level and report the results in table 5. By comparing the results in panel A or table 5 and panel A of table 3, we find that winsorizing increases the contemporaneous correlations of ASBP weighted with number of trades and volume, but not much of ASBP weighted with log(volume), as expected. However, log(volume) weighting still generates the largest correlations. Across columns, *HHW*, *delta_per_share*, *equal_weighting* ASBP are more informative than *one_pair* and *delta_per_dollar* ASBP, consistent with previous results Panel B of table 5 shows that the winsorizing effect is not clear on forecasting excess returns. The correlations slightly decrease in winsorized sample for most aggregation methods but surprisingly *HHW* and volume weighted ASBP, the previously most informative aggregation, gains slightly more correlation, indicating that data smoothing may enhance only the predictability of accurately aggregated option buy pressures. Panel C and D of table 5 show that in the winsorized sample, AVBP has larger correlations to both contemporaneous and future volatilities for all aggregation methods. The order of magnitude of univariate correlations, however, remains the same, i.e. *HHW* and log(volume) weighted AVBP always has the largest correlations. The results presented here clearly show that our conclusion about the effectiveness of aggregation methods is not driven by the extreme observations.

[Table 5 about here.]

4.4. Time series regression

We have shown the importance of aggregation methods in extracting information content from option transactions. In this subsection, we run time series regressions of equation 6 and equation 7 to further test the effectiveness of aggregate option buy pressures in the multivariate setup. For reporting simplicity, we only report the results for our proposed *HHW* and log(volume) weighted aggregate option buy pressures because this aggregation method has shown great consistency and efficiency in previous discussions. Using other aggregating methods generates weaker coefficients in the regression results and the significant coefficients appear at different time lags. However, the general pattern is consistent and will not change our conclusion. In the rest of the paper, we will also stick to this aggregation method for time series analysis if not otherwise specified.

We report the estimation results of equation 6 and equation 7 in table 6. Contemporaneous ASBP is strongly correlated to excess returns with a coefficient of 0.318. ASBP lags, though, are negatively correlated to excess returns. For example, the first lag ASBP has a coefficient of -0.079 while the univariate correlation of lag ASBP and excess returns is significantly positive. In fact, the first five lags are all negative and significant at 1% level. We report up to the tenth lag because longer lags become statistically insignificant. The result demonstrates that the contemporaneous impact of ASBP dominates lagged ASBP. The volatility regression result shows that all AVBP terms have positive and significant correlations. The magnitude of the coefficient does not decrease sharply in lags, suggesting the impact of AVBP lasts longer in the underlying market than ASBP. Again, the results clearly show that option transactions have important information about the underlying market.

[Table 6 about here.]

5. Aggregate option buy pressures under changing market conditions

5.1. Extreme option buy pressures

We have documented that option trading is informative about the underlying market in previous sections. The relationships, however, may not be linear. On one hand, large option trading imbalance may imply significant new information arriving at the market, increasing investors' awareness and hence the correlations of aggregate option buy pressures and the movement in the underlying market. On the other hand, large option trading imbalance is often a result of trades at large sizes. If such trades have been negotiated in the upstairs market before reporting to the exchange, the aggregate option buy pressures will become less informative in those periods. The overall effect is thus unclear. In this subsection, we analyze the tails of the aggregate option buy pressures to see if the underlying market sensitivity responds to the strength of aggregate option buy pressures.

For each aggregate option buy pressure, we construct a subsample of observations with top and bottom 1% values and repeat the same univariate analysis in the previous section. The results are reported in table 7. Panel A shows that the contemporaneous correlation with excess returns almost doubles for all ASBP, indicating that large imbalance in option trading does attract more attention across markets. In fact, the contemporaneous correlations of log(volume) weighted ASBP and excess returns can reach 0.75. Panel B shows that the correlations of ASBP and future excess returns also increase significantly although the statistical significance slightly decrease. For example, *HHW* and volume weighted ASBP has a correlation of 0.062 to future excess returns, compared to 0.011 in the full sample. The results clearly show that the options market becomes more informative about the underlying stock price in the presence of large aggregate stock buy pressures. However, extreme AVBP are no longer significantly correlated with stock volatilities as shown in panel C and D. Although the correlations also increase in magnitude, the only significant contemporaneous correlations are of *HHW*(0.043) and *one_pair*(0.041) AVBP weighted with log(volume) while no AVBP is significantly correlated with future volatilities. It seems that extreme aggregate option buy pressures contribute more in price discovery than in volatility discovery.

[Table 7 about here.]

Alternatively, we run the following time series regressions in the full sample to study the informativeness of extreme aggregate buy pressures:

$$Ret_t = \alpha 0 + \sum_{i=0}^{k} \alpha 1_k ASBP_{t-k} + \sum_{i=0}^{k} \alpha 2_k D1_k + \sum_{i=0}^{k} \alpha 3_k (D1_k * ASBP_{t-k}) + \varepsilon_t, \tag{8}$$

and

$$Vol_t = \beta 0 + \sum_{i=0}^{j} \beta 1_k AVBP_{t-j} + \sum_{i=0}^{j} \beta 2_k D2_k + \sum_{i=0}^{k} \beta 3_k (D2_k * ASBP_{t-k}) + \xi_t, \tag{9}$$

where D1 and D2 are dummy variables equal to 1 for periods with the top or bottom 1% value of ASBP and AVBP respectively, and 0 otherwise. The coefficients of the interaction terms, α3 and β3 will then illustrate the power of aggregate option buy pressures in the tails. For reporting simplicity, we again choose *HHW* and log(volume) weighted aggregate buy pressures and report the results of regression model with two lags in table 8. The surprising result is that $\alpha 3_{t-0}$ is -0.102 with t-value of -22.11, indicating that the contemporaneous correlation between ASBP and excess returns becomes weaker in the tails of ASBP controlled for lag ASBP. However, the predicting power of the tails of ASBP is strong. Both $\alpha 3_{t-1}$ (0.044) and $\alpha 3_{t-2}$ (0.024) are positive and significant at 1% level. The ASBP terms keep the same signs, i.e. the contemporaneous ASBP has a positive coefficient and lag ASBP have negative coefficients. Unlike ASBP, however, AVBP becomes less informative in the tails. $\beta 3_{t-0}$ is negative but insignificant and $\beta 3_{t-1}$ and $\beta 3_{t-2}$ are significantly negative, indicating that the predicting power of AVBP decreases when the options market experiences large volatility buy pressures.

[Table 8 about here.]

5.2. Extreme option trading volume

Holowczak, Simaan, and Wu (2006)) document that the options market becomes more informative when option trading is active. Intuitively, if uninformed trading intention does not correlate with trade occurrence, large trading volume is likely to be due to active informed trading. And we should expect the aggregate option buy pressures to be more informative in such periods. In this subsection, we repeat our univariate analysis in a subsample of observations with top 1% option trading volumes and report the results in table 9. Panel A shows that the contemporaneous correlations of ASBP and excess returns significantly increase during periods when options are actively traded. However, the increment in magnitude is less than in panel A of table 7 except for ASBP weighted with number of trades. The maximum contemporaneous correlation of ASBP and excess returns is around 0.5. The correlations of ASBP and future excess returns also increase although log(volume) weighted ASBP lose forecasting power as shown in panel B. Panel C and D show that the correlations of AVBP and volatilities also increase when the options market is active. However, only log(volume) weighted AVBP remain significant. These results demonstrate that option trading intensity seems to increase only the price information about the underlying stock.

[Table 9 about here.]

We also run the following time series regressions in the full sample:

$$Ret_t = \alpha 0 + \sum_{i=0}^{k} \alpha 1_k ASBP_{t-k} + \sum_{i=0}^{k} \alpha 2_k D_k + \sum_{i=0}^{k} \alpha 3_k (D_k * ASBP_{t-k}) + \varepsilon_t, \quad (10)$$

and

$$Vol_t = \beta 0 + \sum_{i=0}^{j} \beta 1_k AVBP_{t-j} + \sum_{i=0}^{j} \beta 2_k D_k + \sum_{i=0}^{k} \beta 3_k (D_k * ASBP_{t-k}) + \xi_t, \quad (11)$$

where D is a dummy variable equal to 1 when the total option trading volume reaches the top 1%, and 0 otherwise. If the option trading volume is irrelevant to the information content, we would expect the coefficients of the interaction terms, $\alpha 3$ and $\beta 3$ to be insignificant. Table 10 reports the results of these regressions. The coefficient of $ASBP_{t-0} * D_{t-0}$ is -0.152 with t-value of -20.30, indicating that the high trading volume reduces the contemporaneous impact of ASBP. However, The predicting power of ASBP increases as the lag interaction terms have significant coefficients of 0.063 (t-value=8.43) and 0.014 (t-value=1.94). The informativeness of AVBP seems to be uncorrelated with the options volume as none of the interaction terms has significant coefficient.

[Table 10 about here.]

5.3. Shock in aggregate option buy pressures

Other than the level effect shown in the previous discussion, we are also interested in the marginal effect of aggregate option buy pressures. We want to test if the shocks in aggregate option buy pressures can be used as a better indicator for the information content in the options market.

5.3.1. First order difference of buy pressures

To measure the shocks in aggregate option buy pressures, we consider the first order difference of each ASBP and AVBP time series. We repeat the univariate correlation analysis with the first order difference terms instead of level terms and report the results in table 11. The strong statistical correlations in panel A and panel B clearly show that the shocks in ASBP has information about the price movement. However, the shocks in AVBP do not seem to be informative about the volatility movement.

[Table 11 about here.]

We then run the following time series regressions and report results in table12:

$$Ret_t = \alpha 0 + \sum_{i=0}^{k} \alpha 1_k dASBP_{t-k} + \varepsilon_t, \tag{12}$$

and

$$Vol_t = \beta 0 + \sum_{i=0}^{j} \beta 1_k dAVBP_{t-j} + \xi_t, \tag{13}$$

where $dASBP_{t-k}$ and $dAVBP_{t-k}$ denote the first order differences of ASBP and AVBP at lag k, respectively. The results of the return regression are consistent with the univariate analysis. dASBP is informative about excess returns up to 10 minutes later and the magnitude of the coefficients do not decrease sharply with time lags. The coefficient of $dASBP_{t-5}$ is 0.100, one third of the coefficient of the contemporaneous shock, $dASBP_{t-0}$. The volatility regression sheds some light on the impact of shocks in AVBP. Consistent with the univariate analysis, contemporaneous and lag 1 dAVBP are not informative about stock volatilities. However, lag 2 to lag 5 dAVBP have positive and significant coefficients, indicating shocks in AVBP also contain volatility information but the underlying market takes more time to respond.

[Table 12 about here.]

5.3.2. Expected and residual buy pressures

We use an alternative way to measure the expected aggregate option buy pressures and shocks. We perform exponential smoothing for each ASBP and AVBP time series with smoothing parameter of 0.06. The smoothing parameter is randomly chosen at a reasonable value and does not generate conflicting results with alternative smoothing parameters. We use the fitted values as the expected aggregate option buy pressures and the residuals as shocks.

[Table 13 about here.]

Table 13 shows the univariate correlations of expected aggregate buy pressures and the underlying market movement. It is noted that the expected ASBP is negatively correlated with both contemporaneous and one-period ahead excess returns and the expected AVBP is positively correlated with volatilities. The univariate correlations are significant for almost all aggregation methods. Table 14 shows the univariate correlations of shocks in aggregate buy pressures and the underlying market movement. ASBP shocks are positively correlated with excess returns while AVBP shocks are negatively correlated with volatilities. Comparing these two tables, we find that the overall market impact mainly comes from shocks of ASBP and the expected AVBP.

[Table 14 about here.]

We also run the time series regressions on expected and residual aggregate option buy pressures separately and report the results in table 15. Panel A shows that the expected ASBP has a positive and significant contemporaneous coefficient of 0.112. The lag expected ASBP has only negative and significant coefficients at lag 1, 6, and 9. The expected AVBP also has a positive and significant contemporaneous coefficient of 0.099. The lag expected AVBP has only one positive and significant coefficient at lag 10. Panel B shows much of the market impact of ASBP comes from shocks. The contemporaneous ASBP shock has a coefficient of 0.312. Lag ASBP shocks are also significant at lag 1, 2, 3, 7, and 9 although the sign reverses several times. AVBP shocks, on the other hand, are not so informative as the only significant coefficient (-0.003) in the regression belongs to lag 10 AVBP shock.

[Table 15 about here.]

From the analysis in this subsection, we find that shocks in ASBP are quite informative about the stock price but shocks in AVBP do not contain much information about volatilities.

5.4. Intra-day dynamics

The intra-day dynamic is important once we enter the high frequency world. In this subsection, we divide the whole sample into thirteen 30-minute subsamples to examine the intra-day dynamics of links between aggregate option buy pressures and the underlying market. Since we exclude the first 15 and last 5 trading minutes from our sample, our first subsample contains trades in only 15 minutes from 9:45:00 to 9:59:59 and the thirteenth subsample contains trades in 25 minutes from 15:30:00 to 15:54:59. Table16 shows that even with trimmed data at the market open and close, the options market is more active in the beginning and the end than the rest of the day in terms of both number of trades and volume. The first 30-minute subsample has 582 trades per day on average and the total volume reaches an average of 45,258 contracts.

Both number of trades and volume present a clear U-shape pattern along the subsamples with the bottom at the eighth subsample (13:00:00-13:29:59). Toward the end of the day, the options market becomes active again. However, the average trade size before the market close is 54.4 contracts, less than 77.8 contracts in market opening, indicating large trades are more likely to happen in the morning.

Table 16 also reports the univariate correlations between aggregate option buy pressures and excess returns and volatilities. The aggregate option buy pressures are constructed from *HHW* and log(volume) weighting except for the ASBP to forecast excess returns, where volume weighting is used instead of log(volume) because volume weighting has the largest univariate forecasting correlations. Table16 shows the contemporaneous ASBP is always positively and significantly correlated with excess returns and the correlation is the strongest in the market open. However, ASBP cannot predict excess returns in the market open. The predictability then increases gradually and peaks at noon (12:30:00-12:59:59). In the afternoon, ASBP again becomes less informative about future returns but regains predictability in the last half hour before the market close. The correlation between AVBP and volatilities is more complex. The contemporaneous correlation is the strongest in the 4th, 11th and 12th periods. It is also significant in the market open and around noon. However, for the rest of the day including the market close, AVBP does not seem to have contemporaneous impact on stock volatilities. The forecasting power of AVBP becomes significant from 11:30:00 to 12:29:59, and from 13:30:00 to 15:29:59 while for the rest of the day including the market open and close, AVBP cannot predict future volatilities.

[Table 16 about here.]

5.5. Exploring predictability of aggregate option buy pressures

We have focused on both contemporaneous impact and forecasting power of aggregate option buy pressures in previous sections. In this subsection, we further explore the predictability of aggregate option buy pressures by examining different aggregating and forecasting horizons. At every second t in our sample, we construct overlapping aggregate option buy pressures up to k seconds ago (t-k) and examine the correlations between these option buy pressures and the following j seconds (t+j) returns and volatilities. We choose *HHW* and volume weighted ASBP to forecast returns and *HHW* and log(volume) weighted AVBP to forecast volatilities. The results are reported in table 17. Panel A shows that ASBP has greater predicting power in short horizons. t-5 ASBP and t+5 excess returns have the largest correlation of 0.012, statistically significant at 1% level. The correlation decreases in both aggregating horizon and forecasting horizon. Using ASBP aggregated in less than one minute, we are able to predict excess returns up to five minutes. For example, t-10 ASBP has a significant correlation of 0.004 to t+300 returns. On the contrary, AVBP has greater predicting power in long horizons as shown in panel B. The forecasting correlation increases almost monotonically in both aggregating horizon and forecasting horizon. t-300 AVBP has significant correlation of 0.040 to t+120 and t+300 volatilities. In shorter aggregating or forecasting horizons, AVBP still has significant correlations with future volatilities although the magnitude is smaller.

[Table 17 about here.]

6. Conclusion

In this paper, we have shown that option trading contains significant amount of information about the underlying stock market. In particular, aggregate stock buy pressures are positively correlated with both contemporaneous and future excess stock returns, and aggregate volatility buy pressures are positively correlated with both contemporaneous and future volatilities. To efficiently extract such information, however, one needs to rely on good aggregating methods across different option contracts and trades. We summarize our findings from the comparative analysis as follows:

1. Both trade occurrence and size are important considerations when aggregating option order flows for the same option contract. Generally log(volume) weighting is a better choice than number of trades or volume because log(volume) nests the effect of the other two methods. However, when we measure ASBP to forecast excess returns, volume weighting generates more significant results.
2. When aggregating option order flows across strikes and maturities, our proposed *HHW* generates more informative aggregate buy pressures than the other aggregations. Nevertheless, one should still be able to extract significant amount of information about the underlying market using *Greek_per_share* or *equal_weighting* aggregation. *one_pair* and *Greek_per_dollar* are less efficient though as they quite often fail generating informative aggregate option buy pressures.

We also perform conditional time series regressions and subsample analysis with the most efficient aggregation methods. Our results demonstrate that the options market becomes more informative about price discovery when experiencing large trading volume or large ASBP while the volatility information content is insensitive to trading intensity. Shocks in ASBP are more informative about the stock price than expected ASBP while shocks in AVBP do not seem to provide extra information. More importantly, we show that the forecasting power of ASBP decreases in both aggregating and forecasting horizons while the forecasting power of AVBP increases in both aggregating and forecasting horizons. We are able to predict up to five minutes excess returns using ASBP constructed in less than one minute. And we can also predict stock volatilities up to five minutes. These findings confirm our hypothesis of the existence of information trading in the options market. Our results can provide useful guidance for empirical research on the options market microstructure.

References

Anand, A., and S. Chakravarty, 2003, “Stealth Trading in Options Market,” working paper, Syracuse University.

Ané, T., and H. Geman, 2000, “Order Flow, Transaction Clock and Normality of Asset Returns,” *Journal of Finance*, 55(5), 2259–2284.

Anthony, J. H., 1988, “The Interrelation of Stock and Options Market Trading-Volume Data,” *Journal of Finance*, 43, 949–964.

Bhattacharya, M., 1987, “Price Changes of Related Securities: The Case of Call Options and Stocks,” *Journal of Financial and Quantitative Analysis*, 22, 1–15.

Black, F., 1975, “Fact and Fantasy in the Use of Options,” *Financial Analysts Journal*, 31, 36–72.

Black, F., and M. Scholes, 1973, “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81(3), 637–654.

Bollen, N. P., and R. E. Whaley, 2004, “Does Net Buying Pressure affect the Shape of Implied Volatility Functions?,” *Journal of Finance*, 59(2), 711–753.

Cao, C., Z. Chen, and J. M. Griffin, 2005, “Information Content of Option Volume Prior to Takeovers,” *Journal of Business*, 78, 1073–1109.

Carr, P., and L. Wu, 2004, “Static Hedging of Standard Options,” working paper, Bloomberg and Baruch College.

Carr, P., and L. Wu, 2009, “Variance Risk Premiums,” *Review of Financial Studies*, 22(3), 1311–1341.

Chakravarty, S., H. Gulen, and S. Mayhew, 2004, “Informed Trading in Stock and Option Markets,” *Journal of Finance*, 59, 1235–1258.

Chan, K., Y. P. Chung, and W. M. Fong, 2002, “The Informational Role of Stock and Option Volume,” *Review of Financial Studies*, 14, 1049–1075.

Chan, K., Y. P. Chung, and H. Johnson, 1993, “Why Option Prices Lag Stock Prices: A Trading-Based Explanation,” *Journal of Finance*, 48(5), 1957–1967.

Dickey, D. A., and W. A. Fuller, 1979, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

Easley, D., M. O'Hara, and P. S. Srinivas, 1998, "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–465.

Engle, R., 2004, "Risk and Volatility: Econometric Models and Financial Practice," *American Economic Review*, 94(3), 405–420.

Finucane, T. J., 1991, "Put-Call Parity and Expected Returns," *Journal of Financial and Quantitative Analysis*, 26, 445–457.

Hasbrouck, J., 1995, "One Security, many Markets: Determining the Contributions to Price Discovery," *Journal of Finance*, 50(4), 1175–1199.

Holowczak, R., Y. Simaan, and L. Wu, 2006, "Price Discovery in the U.S. Stock and Stock Options Markets: A Portfolio Approach," *Review of Derivatives Research*, 9, 37–65.

Izzeldin, M., 2007, "Trading volume and the number of trades: a comparative study using high frequency data," working paper, Lancaster University Management School.

Jarnecic, E., 1999, "Trading Volume Lead/Lag Relations Between the ASX and ASX Option Market: Implications of Market Microstructure," *Australian Journal of Management*, 24, 77–94.

Jones, C. M., G. Kaul, and M. L. Lipson, 1994, "Information, Trading, and Volatility," *Review of Financial Studies*, 36(1), 127–154.

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

Manaster, S., and R. J. Rendleman Jr., 1982, "Option Prices as Predictors of Equilibrium Stock Prices," *Journal of Finance*, 37, 1043–1057.

Merton, R. C., 1973, "An Intertemporal Asset Pricing Model," *Econometrica*, 41, 867–887.

Ni, S. X., J. Pan, and A. M. Poteshman, 2008, "Volatility Information Trading in Option Market," *Journal of Finance*, 63, 1059–1091.

Pan, J., and A. M. Poteshman, 2006, “The Information in Option Volume for Future Stock Prices,” *Review of Financial Studies*, 19, 871–908.

Simaan, Y., and L. Wu, 2007, “Price Discovery in the U.S. Stock Options Market,” *Journal of Derivatives*, 15(2), 20–38.

Stephan, J. A., and R. E. Whaley, 1990, “Intraday Price Change and Trading Volume Relations in the Stock and Stock Option Markets,” *Journal of Finance*, 45, 191–220.

Table 1

Data Description

The sample includes all option trades expiring in more than 10 calendar days with underlying ticker 'QQQQ' recorded by the Option Price Reporting Authority (OPRA) between 9:45:00 am and 3:54:59 pm EST from 02/01/2006 to 12/29/2006. Percentages are computed based on the number of trades. abs(delta) is the absolute value of Black-Scholes (1973) model implied *delta* computed with 0 interest rate and dividend rate. Near-maturity options are those expiring within 60 calendar days. The trades are classified into buy, sell, and unclassified categories following Lee and Ready (1991) without reporting lag.

Statistics	All options	Calls	Puts
Number of trades	1,087,778	506,948	580,830
Mean daily number of trades	4,709	2,195	2,514
Mean trade size	65	56	73
Std trade size	694	729	662
Median trade size	5	5	6
Mean daily volume	307,170	122,676	184,494
Percentage of near-the-money $0.375 <= abs(delta) <= 0.625$	44.19	44.92	43.55
Percentage of out-of-the-money $abs(delta) < 0.375$	37.23	34.81	39.34
Percentage of near-maturity	79.91	80.25	79.62
Percentage of buy	52.06	49.32	54.46
Percentage of unclassified	0.87	1.03	0.74

Table 2

Underlying returns, volatility, and option buy pressures

This table reports the main variables in the analysis. Panel A reports the mean and standard deviation of dependent variables. Raw returns are calculated as the difference between log prices (NBBO mid quote) at the beginning and the end of each one minute interval. Excess returns are the residual of an MA(1) model of the raw returns. Volatilities are calculated as the annualized standard deviation of second by second NBBO returns within each one minute interval. Panel B and C report the means and standard deviations of the aggregate stock buy pressures and volatility buy pressures as detailed in section 2. The column heads represent the aggregation methods across option strikes and maturities and the row names are the aggregation methods across trade sizes. Standard deviations are in parentheses.

Panel A: Dependent Variables					
Raw return (bp)	excess return (bp)	volatility (%)			
-0.006	-0.002	3.035			
(4.54)	(4.52)	(1.90)			
Panel B: Aggregate Stock Buy Pressure					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.255	-0.676	-0.267	-0.715	-0.183
	(11.43)	(17.55)	(6.01)	(28.15)	(5.05)
volume	0.750	-10.981	-7.451	10.474	-2.380
	(922.50)	(2826.03)	(1101.64)	(2514.73)	(812.85)
log(volume)	-0.133	-0.953	-0.480	-0.440	-0.253
	(6.34)	(19.76)	(10.24)	(14.09)	(6.17)
Panel C: Aggregate Volatility Buy Pressure					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.068	0.636	3.499	3.527	0.191
	(11.36)	(16.31)	(71.64)	(468.23)	(4.73)
volume	4.323	55.834	334.416	421.502	15.907
	(916.33)	(2768.59)	(18168.46)	(38446.62)	(816.65)
log(volume)	0.162	0.898	5.417	5.903	0.275
	(5.72)	(13.07)	(87.27)	(170.95)	(4.29)

Table 3

Aggregate stock buy pressures and excess returns

This table reports the correlations between the underlying excess returns and the aggregate price buy pressures (ASBP). Excess returns are the residual of an MA(1) model of raw returns. The column heads represent the aggregation methods across option strikes and maturities and the row names are the aggregation methods across trade sizes. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.050[a]	0.187[a]	0.285[a]	0.075[a]	0.192[a]
volume	0.063[a]	0.081[a]	0.105[a]	0.062[a]	0.093[a]
log(volume)	0.221[a]	0.396[a]	0.411[a]	0.336[a]	0.388[a]

Panel B: Correlations between ASBP and t+1 excess returns

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.003	0.000	-0.000	-0.001	0.001
volume	0.011[a]	0.008[b]	0.010[a]	0.004	0.011[a]
log(volume)	0.009[a]	0.004	0.003	0.002	0.006[c]

Table 4

Aggregate volatility buy pressures and realized volatilities

This table reports the correlations between underlying volatilities and the aggregate volatility buy pressures (AVBP). Volatilities are calculated as the annualized standard deviation of second by second NBBO returns within each one minute interval. The column heads represent the aggregation methods across option strikes and maturities and the row names are the aggregation methods across trade sizes. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between AVBP and volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.003	0.004	0.009^{a}	-0.000	0.007^{b}
volume	0.006^{c}	0.009^{a}	0.009^{a}	-0.000	0.010^{a}
log(volume)	0.011^{a}	0.014^{a}	0.005	-0.010^{a}	0.018^{a}
Panel B: Correlations between AVBP and t+1 volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.002	0.005	0.006	0.001	0.004
volume	0.006^{c}	0.007^{b}	0.006	-0.000	0.009^{b}
log(volume)	0.010^{a}	0.019^{a}	0.012^{a}	-0.004	0.020^{a}

Table 5

Winsorizing Effect

This table reports the correlations in Table 3 and Tables 4 using winsorized data. All dependent and independent variables are winsorised at 0.1% level. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.067^{a}	0.206^{a}	0.300^{a}	0.104^{a}	0.207^{a}
volume	0.071^{a}	0.109^{a}	0.129^{a}	0.099^{a}	0.118^{a}
log(volume)	0.225^{a}	0.393^{a}	0.410^{a}	0.337^{a}	0.386^{a}
Panel B: Correlations between ASBP and t+1 excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.003	0.001	0.000	0.000	0.002
volume	0.007^{b}	0.010^{a}	0.008^{b}	0.008^{b}	0.012^{a}
log(volume)	0.007^{b}	0.002	0.002	0.002	0.004
Panel C: Contemporaneous correlations between AVBP and volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.000	0.007^{b}	0.011^{a}	0.001	0.008^{b}
volume	0.005	0.013^{a}	0.010^{a}	0.006	0.015^{a}
log(volume)	0.011^{a}	0.015^{a}	0.005	-0.011^{a}	0.018^{a}
Panel D: Correlations between AVBP and t+1 volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.001	0.007^{b}	0.008^{b}	0.002	0.006^{c}
volume	0.006^{c}	0.011^{a}	0.009^{a}	0.004	0.012^{a}
log(volume)	0.012^{a}	0.019^{a}	0.013^{a}	-0.004	0.021^{a}

Table 6

Time Series Regression Results

This table reports the regression results of equation 6 and equation 7. Excess returns are the residuals of MA(1) on one-minute raw returns measured by basis points. Volatilities are the annualized standard deviation of second-by-second returns in each one-minute observation period measured by percentage. The ASBP and AVBP are constructed from *HHW* and log(volume) weighting detailed in section 2.

Dependent: Excess returns			Dependent: Volatilities		
Coefficient of ASBP at lag	Estimate	t-value	Coeffcient of AVBP at lag	Estimate	t-value
0	0.318	129.03	0	0.006	3.66
1	-0.079	-31.11	1	0.007	4.35
2	-0.026	-10.36	2	0.008	5.30
3	-0.019	-7.40	3	0.005	3.12
4	-0.011	-4.21	4	0.005	3.17
5	-0.008	-3.02	5	0.004	2.38
6	-0.003	-1.14	6	0.004	2.83
7	-0.012	-4.74	7	0.004	2.35
8	-0.004	-1.44	8	0.004	2.44
9	0.002	0.67	9	0.004	2.68
10	-0.007	-2.83	10	0.003	2.00

Table 7

Tails of Aggregate Option Buy Pressures and The Underlying Market

This table reports the subsample results for observation periods experiencing extreme option buy pressures (top 1% and bottom 1%). Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.093[a]	0.230[a]	0.397[a]	0.111[a]	0.257[a]
volume	0.191[a]	0.159[a]	0.209[a]	0.133[a]	0.195[a]
log(volume)	0.507[a]	0.749[a]	0.758[a]	0.649[a]	0.731[a]
Panel B: Correlations between ASBP and t+1 excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.029	-0.005	-0.024	-0.015	0.005
volume	0.055[b]	0.047[b]	0.058[b]	0.025	0.062[b]
log(volume)	0.058[b]	0.005	0.007	-0.009	0.012
Panel C: Contemporaneous correlations between AVBP and volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.030	-0.022	-0.002	0.018	-0.011
volume	0.034	0.020	0.020	-0.014	0.017
log(volume)	0.041[c]	0.027	0.013	-0.010	0.043[c]
Panel D: Correlations between AVBP and t+1 volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.026	-0.012	-0.030	0.026	-0.026
volume	0.033	0.020	0.006	-0.013	0.016
log(volume)	0.021	0.026	0.005	-0.028	0.032

Table 8

Tails of Aggregate Option Buy Pressures and The Underlying Market: Regression Results

This table reports the regression results of equation 8 and equation 9. Excess returns are the residuals of MA(1) on one-minute raw returns measured by basis points. Volatilities are the annualized standard deviation of second-by-second returns in each one-minute observation period measured by percentage. The ASBP and AVBP are constructed from *HHW* and log(volume) weighting detailed in section 2.

Dependent: Excess returns			Dependent: Volatilities		
Coefficient	Estimate	t-value	Coefficient	Estimate	t-value
$intercept$	0.046	3.20	$intercept$	2.961	443.34
$ASBP_t$	0.372	107.37	$AVBP_t$	0.006	2.77
$D1_t$	-0.188	-1.85	$D2_t$	1.296	27.81
$ASBP_t * D1_t$	-0.102	-22.11	$AVBP_t * D2_t$	-0.002	-0.73
$ASBP_{t-1}$	-0.108	-30.26	$AVBP_{t-1}$	0.010	5.02
$D1_{t-1}$	0.349	3.40	$D2_{t-1}$	0.745	15.89
$ASBP_{t-1} * D1_{t-1}$	0.044	9.50	$AVBP_{t-1} * D2_{t-1}$	-0.007	-2.20
$ASBP_{t-2}$	-0.050	-14.57	$AVBP_{t-2}$	0.013	6.34
$D1_{t-2}$	0.059	0.58	$D2_{t-2}$	0.726	15.59
$ASBP_{t-2} * D1_{t-2}$	0.024	5.20	$AVBP_{t-2} * D2_{t-2}$	-0.008	-2.68

Table 9

Informativeness of Aggregate Option Buy Pressures and Trading Volume

This table reports the subsample results for periods with top 1% trading volume in the options market. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.170^a	0.437^a	0.515^a	0.268^a	0.407^a
volume	0.186^a	0.144^a	0.205^a	0.113^a	0.192^a
log(volume)	0.326^a	0.505^a	0.529^a	0.438^a	0.498^a

Panel B: Correlations between ASBP and t+1 excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.020	-0.030	-0.039	-0.026	-0.009
volume	0.058^c	0.057^c	0.072^b	0.025	0.072^b
log(volume)	0.035	-0.006	-0.024	-0.012	0.003

Panel C: Contemporaneous correlations between AVBP and volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.009	0.015	0.013	0.030	0.029
volume	0.037	0.006	0.018	-0.029	0.005
log(volume)	0.063^c	0.039	0.034	0.048	0.068^b

Panel D: Correlations between AVBP and t+1 volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.049	0.042	0.039	0.045	0.037
volume	0.046	0.018	0.004	-0.019	0.020
log(volume)	0.065^c	0.082^b	0.054	0.061^c	0.072^b

Table 10

Option Volume and Aggregate Buy Pressures: Regression Results

This table reports the regression results of equation 10 and equation 11. Excess returns are the residuals of MA(1) on one-minute raw returns measured by basis points. Volatilities are the annualized standard deviation of second-by-second returns in each one-minute observation period measured by percentage. The ASBP and AVBP are constructed from *HHW* and log(volume) weighting detailed in section 2.

Dependent: Excess returns			Dependent: Volatilities		
Coefficient	Estimate	t-value	Coefficient	Estimate	t-value
$intercept$	0.046	3.25	$intercept$	3.002	453.16
$ASBP_t$	0.334	129.94	$AVBP_t$	0.005	3.10
D_t	-0.109	-0.76	D_t	0.770	11.57
$ASBP_t * D_t$	-0.152	-20.30	$AVBP_t * D_t$	0.005	1.13
$ASBP_{t-1}$	-0.092	-34.51	$AVBP_{t-1}$	0.007	4.41
D_{t-1}	0.055	0.38	D_{t-1}	0.398	5.96
$ASBP_{t-1} * D_{t-1}$	0.063	8.43	$AVBP_{t-1} * D_{t-1}$	0.001	0.20
$ASBP_{t-2}$	-0.039	-15.30	$AVBP_{t-2}$	0.010	5.99
D_{t-2}	0.269	1.89	D_{t-2}	0.369	5.56
$ASBP_{t-2} * D_{t-2}$	0.014	1.94	$AVBP_{t-2} * D_{t-2}$	-0.004	-0.75

Table 11

First Order Differences of Aggregate Option Buy Pressures

This table reports the correlations between first order differences of aggregate option buy pressures and the underlying excess returns and volatilities. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.048^{a}	0.163^{a}	0.241^{a}	0.070^{a}	0.161^{a}
volume	0.037^{a}	0.054^{a}	0.069^{a}	0.047^{a}	0.064^{a}
log(volume)	0.161^{a}	0.336^{a}	0.347^{a}	0.278^{a}	0.325^{a}
Panel B: Correlations between ASBP and t+1 excess returns					
	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.000	0.004	0.006^{c}	0.001	0.005
volume	0.011^{a}	0.008^{b}	0.009^{a}	0.005	0.010^{a}
log(volume)	0.012^{a}	0.010^{a}	0.010^{a}	0.007^{b}	0.011^{a}
Panel C: Contemporaneous correlations between AVBP and volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.000	-0.000	0.003	-0.001	0.003
volume	0.000	0.001	0.002	-0.000	0.001
log(volume)	0.001	-0.004	-0.005	-0.004	-0.002
Panel D: Correlations between AVBP and t+1 volatilities					
	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.004	-0.004	-0.004	-0.005	-0.003
volume	0.000	0.000	-0.001	-0.000	0.000
log(volume)	-0.005	-0.002	-0.001	-0.003	-0.003

Table 12

First Order Difference of Aggregate Buy Pressures: Regression Results

This table reports the regression results of equation 12 and equation 13. Excess returns are the residuals of MA(1) on one-minute raw returns measured by basis points. Volatilities are the annualized standard deviation of second-by-second returns in each one-minute observation period measured by percentage. The ASBP and AVBP are constructed from *HHW* and log(volume) weighting detailed in section 2.

Dependent: Excess returns			Dependent: Volatilities		
Coefficient of dASBP at lag	Estimate	t-value	Coeffcient of dAVBP at lag	Estimate	t-value
0	0.300	125.12	0	0.000	0.31
1	0.208	72.31	1	0.002	1.08
2	0.169	54.03	2	0.006	2.42
3	0.139	42.30	3	0.006	2.28
4	0.118	34.85	4	0.006	2.26
5	0.100	29.34	5	0.005	1.84
6	0.087	25.75	6	0.004	1.74
7	0.064	19.62	7	0.003	1.37
8	0.050	15.99	8	0.002	1.02
9	0.040	14.15	9	0.001	0.74
10	0.022	9.09	10	-0.001	-0.42

Table 13

Fitted Aggregate Option Buy Pressures

This table reports the correlations in Table 3 and Tables 4 using fitted aggregate option buy pressures. All aggregate option buy pressures are exponentially smoothed (smoothing parameter=0.06). Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.007^b	-0.011^a	-0.015^a	-0.007^b	-0.010^a
volume	0.002	-0.001	-0.004	0.001	-0.000
log(volume)	-0.006	-0.018^a	-0.019^a	-0.016^a	-0.016^a

Panel B: Correlations between ASBP and t+1 excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.007^c	-0.012^a	-0.016^a	-0.008^b	-0.011^a
volume	-0.002	-0.003	-0.006^c	0.001	-0.003
log(volume)	-0.008^b	-0.020^a	-0.020^a	-0.018^a	-0.019^a

Panel C: Contemporaneous correlations between AVBP and volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.005	0.022^a	0.038^a	0.008^b	0.027^a
volume	0.006^c	0.032^a	0.029^a	0.021^a	0.033^a
log(volume)	0.060^a	0.081^a	0.070^a	0.033^a	0.098^a

Panel D: Correlations between AVBP and t+1 volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	0.006^c	0.021^a	0.038^a	0.008^b	0.027^a
volume	0.004	0.032^a	0.028^a	0.022^a	0.033^a
log(volume)	0.060^a	0.080^a	0.069^a	0.034^a	0.097^a

Table 14

Residual Aggregate Option Buy Pressures

This table reports the correlations for the residuals of aggregate option buy pressures after exponentially smoothed (smooth parameter=0.06). Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Contemporaneous correlations between ASBP and excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	0.054^a	0.194^a	0.308^a	0.108^a	0.200^a
volume	0.062^a	0.080^a	0.104^a	0.064^a	0.091^a
log(volume)	0.221^a	0.409^a	0.423^a	0.364^a	0.394^a

Panel B: Correlations between ASBP and t+1 excess returns

	one_pair	equal_weighting	delta_per_share	delta_per_dollar	HHW
number_of_trades	-0.001	0.004	0.005	0.002	0.005
volume	0.011^a	0.009^a	0.011^a	0.005	0.012^a
log(volume)	0.011^a	0.010^a	0.009^a	0.009^b	0.012^a

Panel C: Contemporaneous correlations between AVBP and volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.005	-0.002	-0.001	-0.005	0.001
volume	0.005	0.002	0.004	-0.003	0.004
log(volume)	-0.002	-0.007^c	-0.011^a	-0.015^a	-0.005

Panel D: Correlations between AVBP and t+1 volatilities

	one_pair	equal_weighting	vega_per_share	vega_per_dollar	HHW
number_of_trades	-0.004	-0.001	-0.004	-0.003	-0.002
volume	0.005	0.001	0.001	-0.003	0.002
log(volume)	-0.003	-0.001	-0.005	-0.009^a	-0.003

Table 15
Expected and Residual Aggregate Buy Pressures: Regression Results
This table reports the regression results of equation 6 and equation 7 using the fitted and residual aggregate option buy pressures as independent variables separately. Excess returns are the residuals of MA(1) on one-minute raw returns measured by basis points. Volatilities are the annualized standard deviation of second-by-second returns in each one-minute observation period measured by percentage. The ASBP and AVBP are constructed from *HHW* and log(volume) weighting detailed in section 2.

Panel A: Expected aggregate option buy pressures					
Dependent: Excess returns			Dependent: Volatilities		
Coefficient of ASBP at lag	Estimate	t-value	Coeffcient of AVBP at lag	Estimate	t-value
0	0.112	2.49	0	0.099	3.79
1	-0.151	-2.14	1	0.031	0.84
2	-0.059	-0.83	2	-0.051	-1.39
3	0.049	0.69	3	0.002	0.05
4	0.089	1.26	4	-0.018	-0.50
5	0.005	0.07	5	0.011	0.30
6	-0.199	-2.82	6	-0.011	-0.30
7	0.104	1.48	7	0.000	0.00
8	0.120	1.71	8	0.003	0.08
9	-0.171	-2.45	9	-0.022	-0.61
10	0.075	1.70	10	0.102	3.98

Panel B: Shocks in aggregate option buy pressures					
Dependent: Excess returns			Dependent: Volatilities		
Coefficient of ASBP at lag	Estimate	t-value	Coeffcient of AVBP at lag	Estimate	t-value
0	0.312	128.03	0	0.000	-0.16
1	-0.065	-25.78	1	0.001	0.64
2	-0.016	-6.50	2	0.003	1.78
3	-0.010	-4.12	3	0.000	-0.24
4	-0.003	-1.28	4	0.000	-0.20
5	-0.001	-0.27	5	-0.002	-1.00
6	0.004	1.53	6	-0.001	-0.63
7	-0.005	-2.19	7	-0.002	-1.16
8	0.002	0.98	8	-0.002	-1.22
9	0.008	3.24	9	-0.002	-1.19
10	0.001	0.52	10	-0.003	-2.08

Table 16

Intra-day Dynamics

This table reports the correlations between aggregate option buy pressures and underlying excess returns and volatilities in 30-minute subsamples. Option buy pressures are aggregated with our proposed HHW method across strikes and maturities and with log(volume) weighting across trade sizes with the exception that ASBP is aggregated with volume weighting for forecasting t+1 excess returns. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Subsample	Average Number of Trades	Average Volume	ASBP		AVBP	
			t+0 ret	t+1 ret	t+0 vol	t+1 vol
9:45:00-9:59:59	582	45258	0.424^{a}	0.009	0.030^{c}	0.005
10:00:00-10:29:59	546	42015	0.434^{a}	0.011	-0.010	-0.008
10:30:00-10:59:59	432	29429	0.384^{a}	0.027^{b}	-0.005	0.006
11:00:00-11:29:59	380	24890	0.380^{a}	0.021^{c}	0.036^{a}	0.014
11:30:00-11:59:59	326	19632	0.366^{a}	0.004	0.018	0.028^{b}
12:00:00-12:29:59	311	17372	0.359^{a}	0.020^{c}	0.010	0.030^{b}
12:30:00-12:59:59	296	18047	0.371^{a}	0.049^{a}	0.008	0.001
13:00:00-13:29:59	287	17444	0.365^{a}	-0.024^{b}	0.022^{c}	0.017
13:30:00-13:59:59	299	19028	0.371^{a}	0.030^{b}	0.027^{b}	0.042^{a}
14:00:00-14:29:59	353	28524	0.409^{a}	-0.002	0.002	0.021^{c}
14:30:00-14:59:59	373	21287	0.376^{a}	0.018	0.034^{a}	0.041^{a}
15:00:00-15:29:59	381	23299	0.359^{a}	-0.010	0.042^{a}	0.037^{a}
15:30:00-15:54:59	520	28290	0.372^{a}	0.035^{a}	0.010	0.009

Table 17

Predictability of Aggregate Option Buy Pressures Over Different Horizons

This table reports the correlations between overlapping aggregate option buy pressures and underlying returns and volatilities. Aggregate option buy pressures are measured at each second during sample hours. *HHW* and volume weighted ASBP is used to forecast excess returns and *HHW* and log(volume) weighted AVBP is used to forecast volatilities. Superscription a, b, and c represent significance level at 1%, 5%, and 10%, respectively.

Panel A: Forecasting returns with ASBP

	$t+5ret$	$t+10ret$	$t+30ret$	$t+60ret$	$t+120ret$	$t+300ret$
$t-5asbp$	0.012^a	0.011^a	0.009^a	0.007^a	0.006^a	0.003^a
$t-10asbp$	0.011^a	0.011^a	0.009^a	0.008^a	0.006^a	0.004^a
$t-30asbp$	0.009^a	0.009^a	0.010^a	0.009^a	0.007^a	0.003^a
$t-60asbp$	0.007^a	0.008^a	0.008^a	0.008^a	0.005^a	0.002^a
$t-120asbp$	0.005^a	0.006^a	0.006^a	0.005^a	0.003^a	0.001
$t-300asbp$	0.003^a	0.003^a	0.003^a	0.002^a	0.000	-0.001

Panel B: Forecasting volatilities with AVBP

	$t+5vol$	$t+10vol$	$t+30vol$	$t+60vol$	$t+120vol$	$t+300vol$
$t-5avbp$	0.003^a	0.003^a	0.005^a	0.006^a	0.007^a	0.007^a
$t-10avbp$	0.004^a	0.005^a	0.007^a	0.008^a	0.010^a	0.010^a
$t-30avbp$	0.008^a	0.010^a	0.012^a	0.014^a	0.017^a	0.016^a
$t-60avbp$	0.011^a	0.013^a	0.017^a	0.021^a	0.024^a	0.022^a
$t-120avbp$	0.015^a	0.020^a	0.026^a	0.029^a	0.031^a	0.029^a
$t-300avbp$	0.021^a	0.027^a	0.034^a	0.037^a	0.040^a	0.040^a

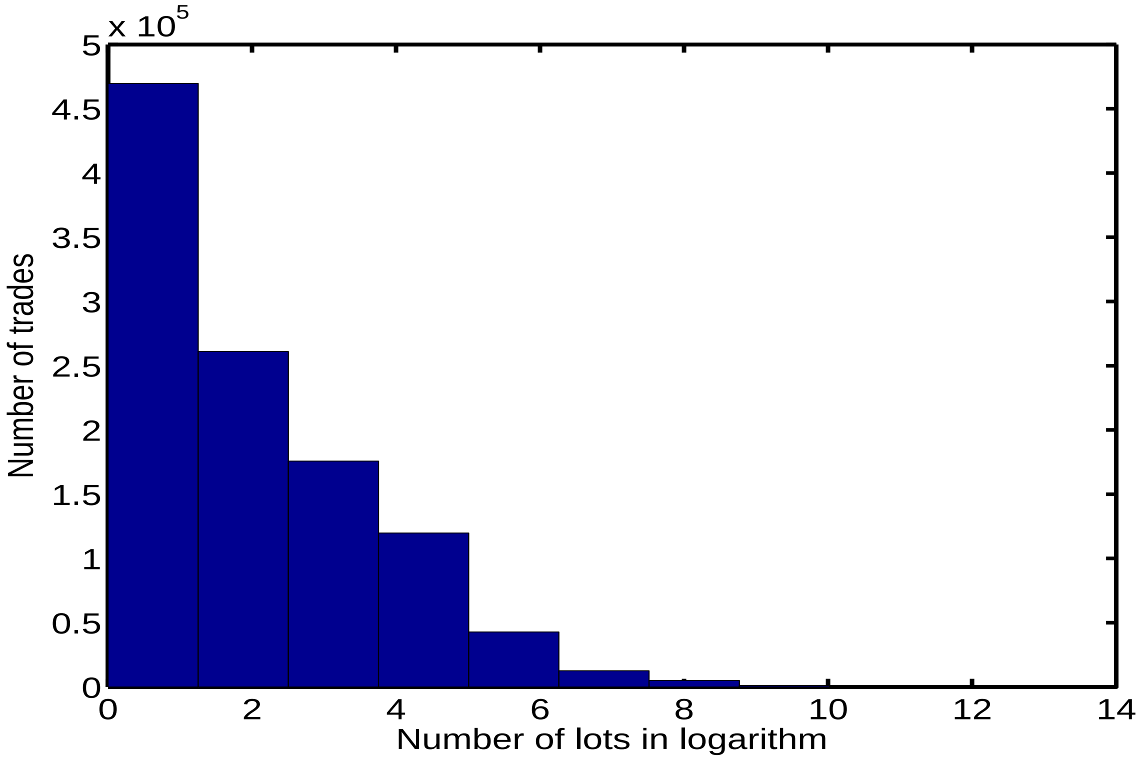
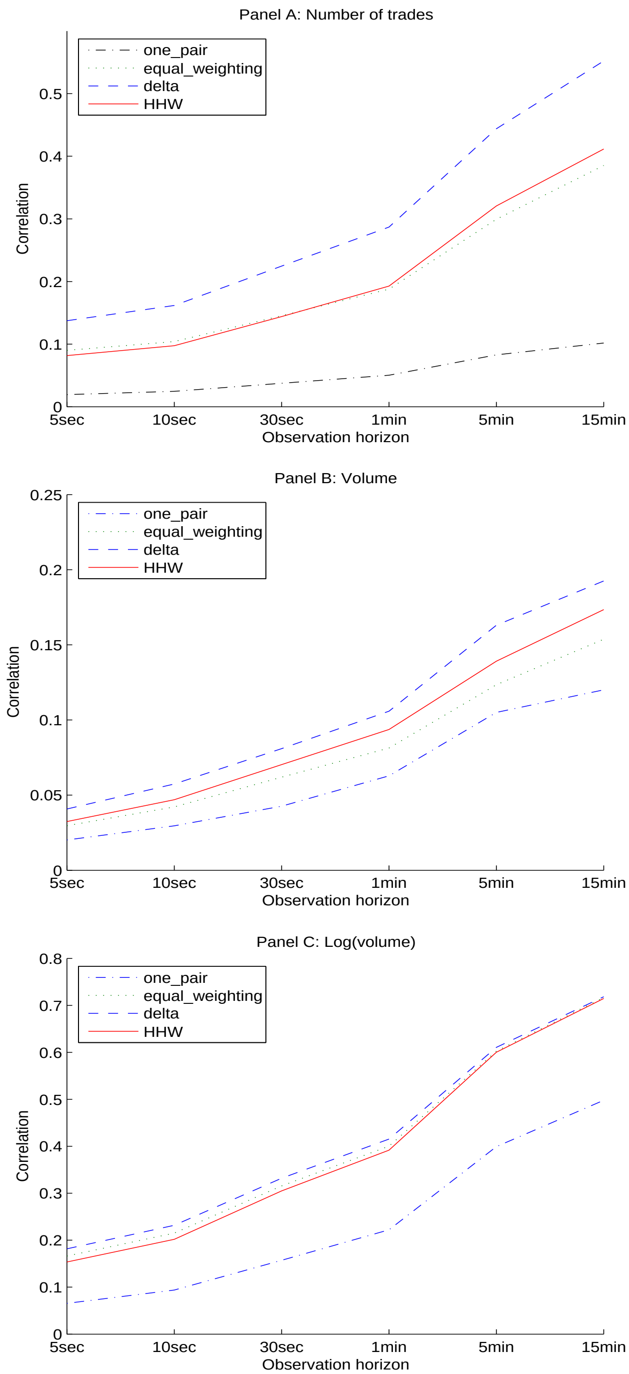


Figure 1
Histogram of QQQQ option transaction size.

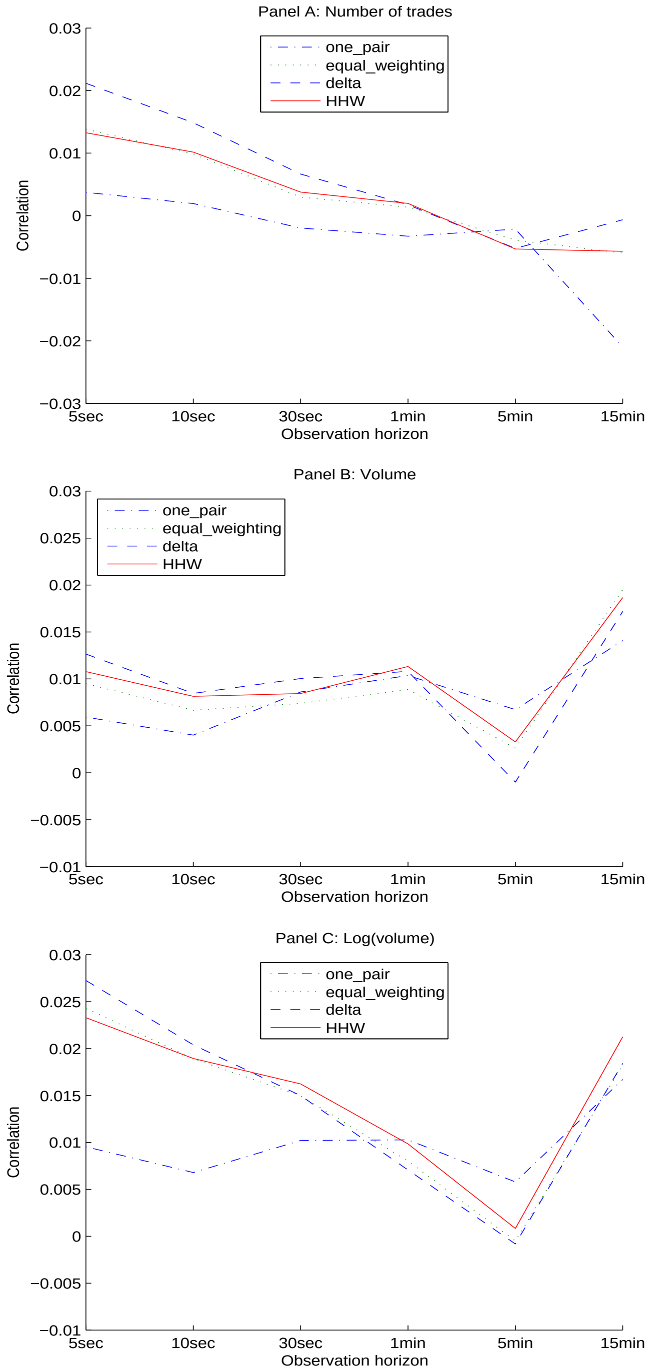
Figure 2

Contemporaneous correlations between ASBP and excess returns over different observation horizons



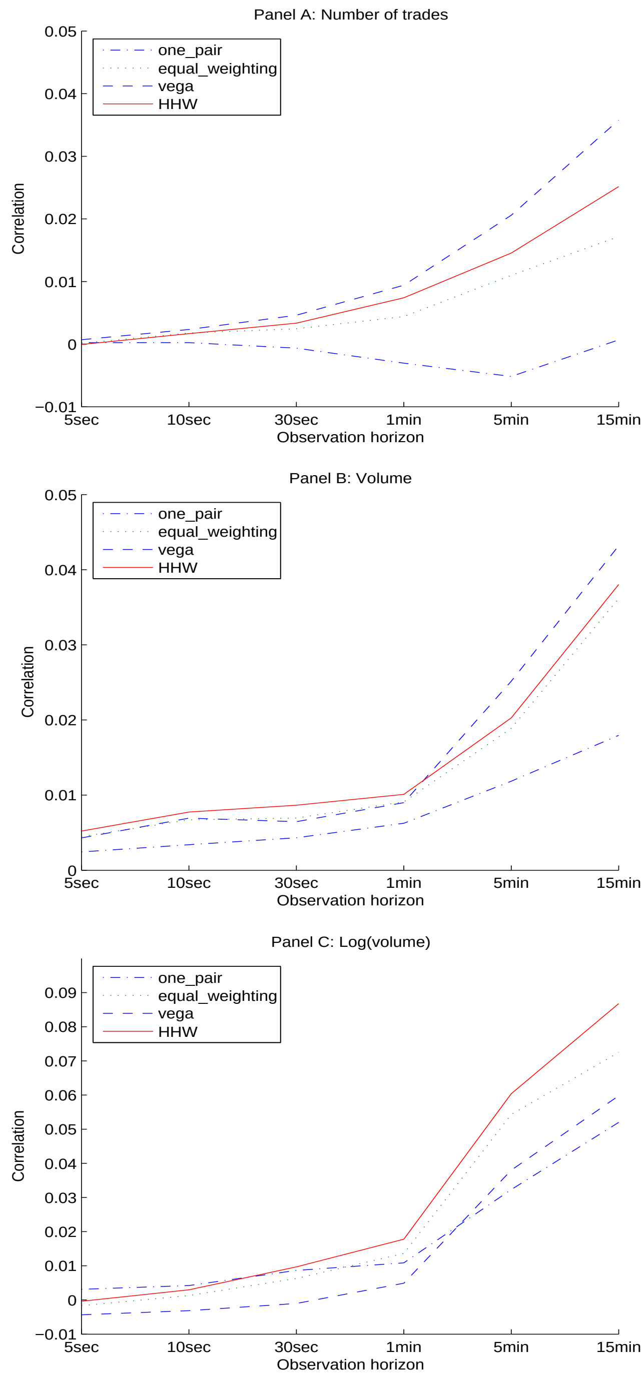
This figure plots the contemporaneous correlations between ASBP and underlying excess returns. The observation window ranges from 5 seconds to 15 minutes.

Figure 3
Forecasting excess returns over different observation horizons



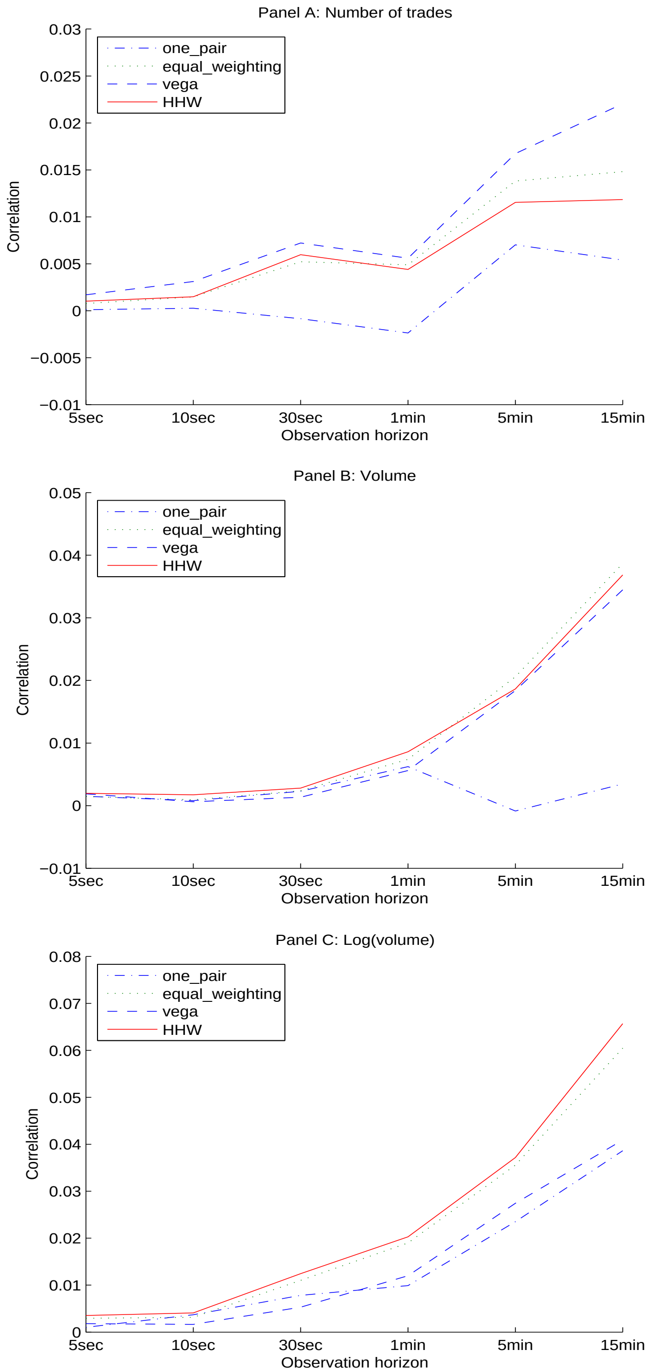
This figure plots the correlations between underlying excess returns and one-period lagged ASBP. The observation window ranges from 5 seconds to 15 minutes.

Figure 4
Contemporaneous correlations between AVBP and volatilities over different observation horizons



This figure plots the contemporaneous correlations between AVBP and underlying volatilities. The observation window ranges from 5 seconds to 15 minutes.

Figure 5
Forecasting volatilities over different observation horizons



This figure plots the correlations between underlying volatilities and one-period lagged AVBP. The observation window ranges from 5 seconds to 15 minutes.