Journal of Financial Economics 107 (2013) 259–283

Contents lists available at SciVerse ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Is there price discovery in equity options? ☆

Dmitriy Muravyev [a,*], Neil D. Pearson [b], John Paul Broussard [c]

[a] Carroll School of Management, Boston College, Chestnut Hill, MA 02467, USA
[b] College of Business, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA
[c] School of Business – Camden, Rutgers University, Camden, NJ 08102, USA

ARTICLE INFO

Article history:
Received 23 October 2011
Received in revised form
23 January 2012
Accepted 2 February 2012
Available online 27 September 2012

JEL classification:
G13
G14

Keywords:
Price discovery
Equity options
Market microstructure
Put-call parity

ABSTRACT

We use tick-by-tick quote data for 39 liquid US stocks and options on them, and we focus on events when the two markets disagree about the stock price in the sense that the option-implied stock price obtained from the put-call parity relation is inconsistent with the actual stock price. Option market quotes adjust to eliminate the disagreement, while the stock market quotes behave normally, as if there were no disagreement. The disagreement events are typically precipitated by stock price movements and display signed option volume in the direction that tends to eliminate the disagreements. These results show that option price quotes do not contain economically significant information about future stock prices beyond what is already reflected in current stock prices, i.e., no economically significant price discovery occurs in the option market. We also find no option market price discovery using a much larger sample of disagreement events based on a weaker definition of a disagreement, which verifies that the findings for the primary sample are not due to unusual or unrepresentative market behavior during the put-call parity violations.

1. Introduction

This paper addresses the fundamental economic question of identifying information flows between the option and stock markets: Do price quotes for equity options contain important directional information about the future level of the underlying stock price that is not yet reflected in stock price quotes? Using a data set that contains more than three years of tick-by-tick trade and quote data for 39 liquid US stocks and options on them, we address this question by focusing on events when the two markets disagree about the stock price in the sense that the bid-ask range of the option-implied stock price quotes obtained from combining the put-call parity relation with bid and ask option price quotes does not overlap with the actual stock bid-ask range. During these events the option market adjusts bid and ask prices to eliminate the disagreement between the option-implied stock price and the actual stock price. In contrast, the disagreement does not affect the stock market. The behavior of stock price changes conditional on a disagreement event cannot be distinguished from the behavior of stock prices in otherwise similar situations without disagreement. These results indicate that when the option price quotes are inconsistent with stock price quotes, the option quotes do not contain any economically significant information that has not already been reflected in the stock market. In this sense, option price quotes do not participate in the price discovery process for the underlying stock price.

We obtain these results using an approach that involves two main components. First, we identify disagreement

☆ We gratefully thank Nanex and Eric Hunsader for providing the trade and quote data for the options and their underlying stocks, and we thank seminar participants at Texas A&M University, the University of Illinois, the University of Notre Dame, Hong Kong University of Science and Technology, City University of Hong Kong, Hong Kong University, Chinese University of Hong Kong, and especially Greg Bauer, Joel Hasbrouck, Tim Johnson, Stewart Mayhew, Mathew Spiegel, and the referee Hendrick Bessembinder for helpful comments. Dmitriy Muravyev acknowledges financial support from the Irwin Fellowship at the University of Illinois.

* Corresponding author.
E-mail address: muravyev@bc.edu (D. Muravyev).

0304-405X/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jfineco.2012.09.003

events in which the stock prices implied by the option price quotes are inconsistent with the actual stock price quotes. In our primary sample of 81,024 disagreement events we use the put-call parity relation to compute option-implied stock prices, explicitly taking account of the large bid-ask spread in the option market to compute option-implied bid and ask prices for the stock. We identify disagreement events only when the actual bid-ask range of the stock does not overlap with the option-implied bid-ask range, and we require that the difference between the closest points of the bid-ask ranges equals or exceeds a threshold. Thus, the price disagreement events in our primary sample are violations of the law of one price. Violations of the law of one price have received attention in other contexts because they are viewed as providing particularly useful information about the functioning of financial markets.[1] Setting aside execution risk and other impediments to arbitrage, the price disagreement events in our primary sample are potentially profitable opportunities for exchange members. These violations are interesting and informative because they are cases in which it is clear that the option and stock markets disagree about the value of the underlying stock and, thus, are important cases in which one can be confident that trading activity and price changes contain information about the markets' responses to the disagreement.

Violations of the law of one price are also useful in assessing the economic significance of price discovery in the options market. If one finds (as we do) that no economically significant option market price discovery occurs during violations of the law of one price, then it seems unlikely that smaller disagreements with the option market will cause economically significant changes in stock price quotes. We confirm this by studying a much larger sample constructed using a broader definition of a disagreement, and we find that in the larger sample stock price quote changes during disagreements also are not different from their changes in otherwise similar situations that do not involve disagreements. These results confirm that the results from the primary sample are not an artifact created by unusual or unrepresentative market behavior during violations of the law of one price. The larger sample includes a total of almost 4.5 million events, about 55 times as many as in the primary sample. Disagreements in this larger sample occur on average more than 130 times per stock, per day, indicating that price disagreements between the option and stock markets are common.

Second, we compare the changes in both the actual and option-implied stock price quotes in the treatment sample of price disagreement events to the quote changes in a matched sample of otherwise similar observations for which no disagreement exists about the underlying stock prices. The use of the matched sample is necessary because stock and option price quotes can change for reasons other than the price disagreement. Using a classical average treatment effect analysis and also examining average effects conditioned on control variables, we find that option-implied stock prices move toward the actual stock prices, but actual stock prices do not move toward the option-implied prices. Instead, the behavior of the actual stock prices when a disagreement exists cannot be distinguished from their behavior in the absence of a disagreement. Examination of histograms summarizing the distributions of changes in option-implied and actual stock price changes reveals that the disagreement events involve a shift in the distribution of option-implied quotes and, essentially, no alteration in the distribution of actual stock prices. We obtain similar results using the much larger sample constructed using a broader definition of a disagreement.

Consistent with our findings that option quotes either agree with or lag stock price quotes, the violations of the law of one price are typically precipitated by stock price moves. This is inconsistent with the potential explanation that the violations of the law of one price are created by anomalous changes in option price quotes, which then revert to their previous level. In addition, the disagreement events often exhibit signed option volume providing pressure in the direction that will tend to move option price quotes to eliminate the disagreement, but no unusual signed volume in the stock market. This is consistent with stock prices leading option prices, with the stock price being right and the option price being wrong. This finding regarding signed option volume during the disagreement events is also evidence that the events are not an artifact of data problems. In the larger sample based on a broader definition of a disagreement, we also find signed volume in the direction that tends to move the option price quotes toward the stock price quotes, but the average signed volume is not as large. This is to be expected because the disagreements in the broader sample are not clear violations of the law of one price.

Our approach is complementary to the large literature following Hasbrouck (1995) that uses the Hasbrouck "information share" and other measures based on vector error correction models (VECMs). This literature has generally focused on summary statistical measures that indicate the proportions of price discovery that occur in different markets, e.g., the Hasbrouck information share and the "component share" proposed by Harris, McInish, and Wood (2002a, 2002b). Chakravarty, Gulen, and Mayhew (2004) is a notable contribution to this literature, finding that the information share of the option market is 17% using data from 1988 to 1992. Papers using more recent data include Holowczak, Simaan, and Wu (2007) and Dong and Sinha (2012), who find option market information shares that are smaller and larger, respectively, than those found by Chakravarty, Gulen, and Mayhew (2004). Our findings about the magnitudes of the quote changes during disagreement events provide information about the economic significance of the statistical predictability found in the literature using VECMs and, therefore, complements this literature. In a different context, a long tradition exists in that part of the empirical asset pricing literature devoted to predictability and market efficiency that emphasizes the importance

[1] See, e.g., Lamont and Thaler (2003b) for an important contribution to this literature and Lamont and Thaler (2003a) for a survey. In the option market, Ofek, Richardson, and Whitelaw (2004), Battalio and Schultz (2006), and Cremers and Weinbaum (2010) have studied put-call parity violations.

of considering the economic as well as the statistical significance of findings of predictability (see, e.g., Jensen, 1978; Fama, 1991).

We also compute the option market information share for our sample and find that the option market information share is consistently in the single digits, i.e., less than 10%. This evidence of limited option market price discovery using a statistical metric, the Hasbrouck information share, seems consistent with our finding of no economically significant option market price discovery, as it is not unusual for finance researchers to find financial market predictability that is statistically but not economically significant. The single-digit estimates of option market information share also indicate that option market price discovery appears to have declined since the sample period used in some of the earlier literature, e.g., Chakravarty, Gulen, and Mayhew (2004), and, thus, help reconcile our results with those in the earlier literature.

In interpreting our results within the context of the existing literature, it is important to distinguish between the question of whether some option trades are based on information about underlying stock prices and the question of whether option quotes contain information about underlying stock prices in addition to that already reflected in the stock price quotes. There are reasons to think that some option trades are executed by informed investors. Black (1975) argues that the embedded leverage and lack of short-sale restrictions make options attractive for traders with information about future stock prices. Easley, O'Hara, and Srinivas (1998), using a sequential trade model, show that a pooling equilibrium with informed trading in both markets can occur. Meanwhile, many of the market participants who are most likely to possess valuable information about future stock prices, e.g. hedge fund managers, have access to actual leverage and, thus, do not need the synthetic leverage available via options.

On the empirical side, Amin and Lee (1997) and Cao, Chen, and Griffin (2005) find evidence of some informed trading in the options market prior to earnings and takeover announcements, respectively. The most convincing evidence comes from Pan and Poteshman (2006), who show that stocks with low put-call ratios computed from the option trades that open new option positions outperform stocks with high put-call ratios computed from opening trades by more than 40 basis points on the next day.[2] (The magnitudes are smaller if only public information on signed volume is used to compute the put-call ratios.) Recent work by Chan, Kot, and Ni (2010) finds that option opening trades contain information about future stock returns out to horizons of several months.

Our finding that options quotes do not contain incremental information relative to stock quotes might seem surprising because it implies that option quotes are not adjusted immediately in response to possibly informative option trades. But a large literature (e.g., Dechow, Hutton, Meulbroek, and Sloan, 2001; Jones and Lamont, 2002; Desai, Ramesh, Thiagarajan, and Balachandran, 2002; Asquith, Pathak, and Ritter, 2005; Nagel, 2005;Boehmer, Jones and Zhang, 2008; Diether, Lee, and Werner, 2009; Boehmer, Huszar and Jordan, 2010; and others) provides evidence that short interest predicts stock returns out to horizons of months, i.e., that stock quotes are not adjusted immediately to reflect the information in some stock trades. Chan, Kot, and Ni (2010) and Blau and Wade (forthcoming) find that option trades contain less information about future stock returns than short selling ratios. Seen in this light, our findings that option quotes do not contain incremental information relative to stock quotes while some option trades appear to contain information is not puzzling, or at least is less puzzling than the well-established finding that stock quotes are not adjusted immediately in response to some stock trades.

During our sample the overwhelming bulk of option trading was electronic, with market makers generally using auto-quoting algorithms and quotes and trades disseminated almost instantly to participants in both the option and equity markets. In contrast to the previous option market structure in which trading occurred on exchange floors, in the current market structure an option market maker on the exchange where trade occurs does not have any informational advantage relative to other market participants, including market makers on the equity exchanges. This helps explain our findings that option quotes do not contain information not already reflected in stock quotes.

The remainder of this paper is organized as follows. The next section briefly reviews the related literature. Section 3 describes the novel data employed in the paper, and Section 4 describes how we define and identify the disagreement events and the construction of the matched control sample, and it then presents some relevant summary statistics about the disagreement events. Section 5 contains the main results of the empirical analysis of quote changes. Section 6 presents results showing that the main results are also found in various subsets of the data, which serve to show that the results are robust and not an artifact stemming from possible data problems. Section 7 reports the estimates of the Hasbrouck (1995) option market information shares for our sample, and Section 8 briefly concludes.

2. Literature review

Our results are most related to the long-standing literature on the lead–lag relation between stock and options markets.

The evidence on whether option price quotes react to potentially informed trading is mixed. On the one hand, Amin and Lee (1997) show that option bid-ask spreads do not increase during periods of plausibly high information

[2] In addition, Easley, O'Hara, and Srinivas (1998) present evidence indicating that options trading volume leads stock trading volume. Enforcement actions by the US Securities and Exchange Commission provide evidence that at least some investors occasionally use options for illegal informed trading. This evidence, as well as that in Amin and Lee (1997), Cao, Chen, and Griffin (2005), and Pan and Poteshman (2006), uses data from sample periods prior to the structural changes in option trading.

asymmetry, such as prior to earnings announcements, and Vijh (1990) points out that option price quotes do not change after large trades. On the other hand, Chan, Chung, and Fong (2002) extend the approach of Hasbrouck (1991) to multiple markets and find that new information in the options market comes in the form of quote revisions, not trades.

Research on the lead–lag behavior between stock and options markets was started by Manaster and Rendleman (1982), who used end-of-day trade prices to show that option prices lead stock prices. However, during that paper's sample period, options trading continued for 10 minutes after the stock market close, raising the possibility that the result stems from the asynchroneity in the quotes. Bhattacharya (1987), who examines the findings of Manaster and Rendleman (1982), does not rely on a specific econometric model and is the closest study to ours. Using 15-minute quote snapshots for 32 stocks and options on them, he simulates a trading strategy in which stock is bought (sold) when it is quoted below (above) the option-implied stock price by at least a threshold amount. The position is held open for 15 minutes and then closed. The average profitability from this strategy is close to zero and becomes negative after bid-ask spread costs are considered, providing no evidence that options lead stocks. However, a critical aspect of Bhattacharya's test design is that it can detect only whether the option market leads the stock market and not vice versa.

Subsequent literature focused on refining the econometric methodology, starting with simple causality tests by Anthony (1988) and evolving to the VECMs estimated by Chan, Chung, and Fong (2002) and Chakravarty, Gulen, and Mayhew (2004). Other related literature includes Stephan and Whaley (1990), Chan, Chung, and Johnson (1993), De Jong and Donders (1998), Finucane (1999), Diltz and Kim (2005), O'Connor (2005). A recent literature review is provided by Ansi and Ben Ouda (2009). Finucane (1991) was the first to apply call-put parity to study lead–lag behavior between the option and stock markets.

Chakravarty, Gulen, and Mayhew (2004) compute option-implied stock prices from call option prices using the binomial model and lagged implied volatilities. Estimating a VECM for the actual and option-implied stock prices with daily data from 1988 to 1992, they find that the Hasbrouck (1995) "information share" of option quotes is 17%. Chakravarty, Gulen, and Mayhew (2004) and the other papers mentioned above use samples that pre-date the transformation in option market structure that began in 2000. Thus, their findings might not be relevant to the current market.

Holowczak, Simaan, and Wu (2007) use data from 52 trading days during the spring and summer of 2002, the put-call parity relation, and bid-ask midpoints to compute the option-implied stock price and then estimate the Hasbrouck (1995) information shares for a sample of stocks and options on them. They obtain somewhat smaller estimates of the information share of option quotes than did Chakravarty, Gulen, and Mayhew (2004), which they attribute to the change in the option market structure that began in 2000. Additional results in Holowczak, Simaan, and Wu (2007) show that the options quotes become more informative when there are large price movements, large numbers of options trades, or significant signed order flow in the options market. In contrast to most of the price discovery literature, Dong and Sinha (2012) estimate the information shares using data on trades in the Dow 30 stocks and options on them from 2003 to 2009 instead of quotes, and they use 5-minute intervals instead of the 1-second intervals used by others. Due to these methodological differences, their results are not comparable to the others in the literature. Dong and Sinha's estimates of the option-market information share are somewhat greater than those in Chakravarty, Gulen, and Mayhew (2004). They find that the option market information share was greater prior to the arrival of news, and they also present evidence that the option market information share of financial stocks increased during the short-sale ban of 2008.

3. Data description

The primary data used in this research are tick-by-tick trade and quote data for 36 liquid US stocks and three exchange- traded funds (ETFs) along with their options from April 17, 2003 to October 18, 2006, a total of 882 trading days.[3] This is a larger and more recent sample than used in previous studies of the lead–lag relation. The data were obtained from Nanex, which provides real-time option and stock price data to its customers via its NxCore product. The data were archived by Nanex as they arrived from the exchanges at Nanex's server and time-stamped by Nanex to 25 millisecond precision as they arrived. The data come from all US exchanges where a given contract is traded. For trades, transaction price, size, exchange code, and some other information are available. For quotes, exchange-level best quotes and volumes are available. That is, the data include each instance when any exchange adjusts its best quote or quoted volume, even if it this change does not change the national best bid and offer (NBBO). The main reason for limiting the sample size to 39 stocks and ETFs was a data storage limitation.

The sample stocks were those with the highest option market volume during March 2003. The underlying stocks are listed in Table 1. They consist of a number of stocks, e.g., large-capitalization technology stocks, that have persistently high option volume, along with a few smaller-capitalization stocks that happened to be of trading interest during the spring of 2003. The sample stocks cover a significant share of total volume in the options market. Some of the stocks merged, e.g., America Online, or changed a ticker, e.g., Morgan Stanley. For these stocks, data are available only up to the date of the ticker change. This resulted in five stocks (AOL, CPN, MWD, SBC, and NXTL) dropping from the sample prior to the end date. In addition, the QQQ ticker for the Nasdaq

[3] Data are not available for several days during the period, so the total of 882 days does not match the number of trading days from April 17, 2003 to October 18, 2006, inclusive.

Table 1
Summary statistics for the 36 stocks and three ETFs in the primary sample. The median duration is the time (s) from when a disagreement event is triggered to the first time the actual stock quotes return inside the option-implied quotes. The implied spread is the difference between the option-implied ask and bid quotes, each computed using the call-put parity relations Eqs. (1) and (2), in cents. It is computed from the option quotes prevailing at the moment a disagreement event is triggered. The pre-event returns are computed from stock quote midpoints prevailing when each event is triggered and 2 minutes or 10 seconds before the event and are reported in percentage points. The averages in the last row are equal-weighted averages of the stock/ETF averages. Ticker symbols indicated with ⁎ dropped before the end of the sample period.

	{P > IP}-type disagreements					{IP > P}-type disagreements				
Ticker	Number of events	Median duration (seconds)	Mean implied spread (cents)	Two-minute pre-event return (percent)	Ten-second pre-event return (percent)	Number of events	Median duration (seconds)	Mean implied spread (cents)	Two-minute pre-event return (percent)	Ten-second pre-event return (percent)
AIG	845	10.1	12.9	0.18	0.11	838	9.3	13.0	−0.17	−0.11
AMAT	544	13.4	10.1	0.26	0.13	688	16.2	10.1	−0.30	−0.14
AMGN	1,862	5.9	12.1	0.18	0.12	2075	6.4	12.1	−0.18	−0.12
AMR	407	18.4	10.7	0.47	0.22	443	17.5	10.7	−0.38	−0.19
AMZN	4,206	5.1	11.7	0.23	0.15	4223	5.3	11.7	−0.22	−0.15
AOL*	29	38.8	10.5	0.41	0.28	34	42.7	10.0	−0.20	−0.08
BMY	137	14.9	10.7	0.19	0.11	159	30.8	10.4	−0.18	−0.08
BRCM	3,033	3.9	11.4	0.26	0.17	3260	4.3	11.5	−0.27	−0.18
C	526	21.6	10.3	0.14	0.07	539	21.9	10.4	−0.14	−0.07
COF	837	10.7	14.1	0.21	0.12	740	9.0	14.5	−0.20	−0.12
CPN*	27	53.0	10.0	0.48	0.18	41	73.0	10.0	−0.66	−0.26
CSCO	474	14.9	10.0	0.24	0.13	603	15.0	10.0	−0.25	−0.13
DELL	1,518	7.7	10.1	0.21	0.11	1,460	8.5	10.1	−0.20	−0.12
DIA	266	19.0	20.3	0.07	0.03	231	18.1	19.2	−0.10	−0.04
EBAY	4,634	6.0	12.8	0.19	0.12	5241	5.4	12.8	−0.20	−0.13
EMC	126	30.5	10.2	0.28	0.14	179	37.3	10.4	−0.28	−0.12
F	44	34.3	10.1	0.26	0.13	66	40.7	10.6	−0.24	−0.11
GE	255	25.7	10.0	0.14	0.07	279	29.1	10.3	−0.13	−0.06
GM	1,018	14.0	11.1	0.23	0.13	937	14.3	11.0	−0.20	−0.11
HD	420	12.6	10.3	0.17	0.10	510	12.7	10.4	−0.20	−0.09
IBM	791	14.3	12.4	0.12	0.07	843	14.9	12.5	−0.14	−0.07
INTC	1,190	20.3	10.0	0.21	0.11	1419	20.1	10.0	−0.23	−0.11
JPM	410	19.3	10.2	0.15	0.07	425	21.8	10.4	−0.16	−0.08
KLAC	2,619	5.4	12.2	0.21	0.14	2678	4.5	12.2	−0.22	−0.15
MMM	357	11.7	13.1	0.15	0.09	384	9.1	13.1	−0.14	−0.09
MO	490	12.3	12.1	0.15	0.09	497	16.1	11.9	−0.14	−0.09
MSFT	719	17.6	10.0	0.19	0.10	849	20.5	10.0	−0.19	−0.10
MWD*	317	10.6	12.2	0.18	0.11	268	12.8	12.3	−0.17	−0.09
NXTL*	562	12.9	10.2	0.26	0.17	692	12.0	10.2	−0.29	−0.17
ORCL	141	23.2	10.1	0.32	0.18	155	18.1	10.0	−0.34	−0.17
PFE	488	17.9	10.4	0.18	0.10	579	23.8	10.2	−0.16	−0.08
QCOM	3,277	4.3	11.2	0.22	0.14	3,796	4.2	11.1	−0.22	−0.14
QLGC	1,887	5.4	12.2	0.26	0.18	1940	5.6	12.2	−0.26	−0.18
QQQ*	916	30.0	10.0	0.13	0.05	1021	33.4	10.0	−0.13	−0.05
QQQQ	585	5.0	10.1	0.15	0.05	667	4.9	10.1	−0.16	−0.08
SBC*	61	31.7	10.2	0.24	0.11	53	36.8	10.0	−0.20	−0.09
SMH	992	16.8	10.5	0.16	0.07	1000	16.5	10.5	−0.19	−0.08
TYC	301	19.4	10.3	0.23	0.12	325	23.1	10.2	−0.19	−0.10
XLNX	1,375	5.9	10.9	0.28	0.17	1576	5.4	11.0	−0.27	−0.18
XOM	293	10.3	11.4	0.15	0.09	332	10.0	10.9	−0.15	−0.07
Mean or sum	38,979	16.4	11.2	0.22	0.12	42,045	18.3	11.2	−0.22	−0.11

100 ETF dropped from the sample when it was replaced by the QQQQ ticker.

An advantage of a sample that consists of stocks with the highest option trading volume during March 2003 is that these stocks are likely to be those in which the impediments to option market price discovery, e.g., limited options market liquidity, are least important. The sample, however, is not representative of the broader equity markets, and the results might be different for other stocks. For example, it might be that information asymmetries are unimportant for the stocks with the highest option trading volume and that different results would be obtained for a different set of stocks. In evaluating these arguments, an important piece of evidence is that the sample includes a few stocks that happened to have high option trading volume during March 2003 due to corporate events or other corporate news but did not generally have very high option trading volume throughout the sample period. Table A1 of the Internet Appendix shows that the main results are found through the entire sample, including in the stocks that had the lowest average option trading volume during the periods when they were in the sample.[4]

Dividend and split data are taken from the Ivy DB database available from Option Metrics LLC. The Ivy DB database includes dividend amounts as well as declaration, ex-dividend, and payment dates. The risk-free rates are also from Option Metrics, which compute them from London Interbank Offered Rate (LIBOR) quotes and LIBOR forward rates estimated from the prices of Eurodollar futures contracts.

Some analyses use estimates of signed volume for the options or their underlying stocks, and some use estimates of the delta-equivalent signed volume of the options transactions. For stocks, the trade direction is inferred using the Lee and Ready (1991) algorithm. For options, trade direction is inferred by first applying the quote rule to the NBBO. If the trade is at the midpoint of the NBBO, the quote rule is applied to the best bid offer (BBO) from the exchange at which the trade occurs. Based on 6 months of data from 1995, Savickas and Wilson (2003) show that for options the quote rule works better than other common algorithms. Also, unlike stocks, in our data approximately 80% of the options transactions occur at the best bid or ask, which suggests that the quote rule will provide reasonable results for the options. Options delta, needed to estimate delta-equivalent signed volume, is computed using the Black-Scholes-Merton formula and the previous days' implied volatility estimates from Option Metrics.

Earnings announcement data are needed for one analysis. Earnings announcement dates are taken from First Call, and the times of the earnings announcements are hand-checked using LexisNexis.

4. The disagreement and matched control samples

The primary sample consists of disagreement events in which the stock and options markets disagree about the stock price in the sense that the option-implied stock price quotes obtained from the put-call parity relation are inconsistent with the actual stock price quotes. We compare the movement of the actual and option-implied stock price quotes during disagreement events with their movement during otherwise similar control events during which there is no disagreement about the underlying stock prices. Differences in the movement of the quotes between the disagreement and matched control events allow us to draw inferences about the markets in which economically significant price discovery occurs.

4.1. Price disagreement events

Let S_t be the stock price at time t; $C_t(K,T)$ and $P_t(K,T)$ be the prices of call and put options, respectively, with strike K and expiration date T; $\mathrm{PV}_t(D(t,T))$ be the present value as of time t of the dividends with ex-dividend dates falling between the current date t and the option expiration T; and $r(t,T)$ be the interest rate for the period from t to T. We must first deal with the fact that US-traded options on individual equities have American, not European-style, exercise. For this reason, the European put-call parity relation

$$S_t = C_t(K,T) - P_t(K,T) + \mathrm{PV}_t(D(t,T)) + Ke^{-r(t,T)(T-t)} \tag{1}$$

will not be exactly satisfied. Instead, for American options we expect a relation of the form

$$S_t + \nu_t(K,T) = C_t(K,T) - P_t(K,T) + \mathrm{PV}_t(D(t,T)) + Ke^{-r(t,T)(T-t)}, \tag{2}$$

where $\nu_t(K,T)$ is the difference between the early exercise premium of the American call and the early exercise premium of the American put. The term $\nu_t(K,T)$ also absorbs any differences that might arise because the dividend estimates or LIBOR-based interest rates we obtain from Option Metrics are different from those implicit in market prices. For each date and call-put pair defined by K and T, we exploit the high-frequency data to estimate $\nu_t(K,T)$ in the following model-free way.

During each day, for each call-put pair, at every quote update (either bid or ask) for either the call, the put, or the stock, we use quote midpoints to estimate the error from the European put-call parity relation as

$$\varepsilon_j = C_t(K,T) - P_t(K,T) + \mathrm{PV}_t(D(t,T)) + Ke^{-r(t,T)(T-t)} - S_t. \tag{3}$$

The estimate of $\nu_t(K,T)$ used on day t is then the average of the ε_j, that is,

$$\nu_t(K,T) = \sum_{j=1}^{N} \varepsilon_j, \tag{4}$$

where N is the number of option quote updates on day t. This approach assumes that the difference $\nu_t(K,T)$ between the call and put early exercise premia is constant during each day. Given the estimate of the difference $\nu_t(K,T)$, the term $C_t(K,T) - P_t(K,T) - \nu_t(K,T)$ is an estimate of the difference between European option prices and can be used in the European put-call parity formula in place of the difference between European option prices.[5]

Doing this, for the primary sample we define a disagreement event by first constructing option-implied bid and ask prices for the underlying stock using the adjusted option prices and the put-call parity relations

$$\text{Implied Bid}_t(K,T) = C_t^{bid}(K,T) - P_t^{ask}(K,T) - \nu_t(K,T) + \mathrm{PV}_t(D(t,T)) + Ke^{-r(t,T)(T-t)} \tag{5}$$

and

$$\text{Implied Ask}_t(K,T) = C_t^{ask}(K,T) - P_t^{bid}(K,T) - \nu_t(K,T) + \mathrm{PV}_t(D(t,T)) + Ke^{-r(t,T)(T-t)}, \tag{6}$$

where Implied Bid_t and Implied Ask_t are the option-implied bid and ask quotes of the underlying stock and

[4] The five stocks with the lowest average daily option trading volume when they were in the sample were Capital One Financial, Calpine, Ford, SBC Communications, and Xilinx, with ticker symbols COF, CPN, F, SBC, and XLNX, respectively.

[5] This adjustment for the early exercise premia is not crucial. An earlier version of this paper constructed the sample of disagreements using the European put-call parity relation without any adjustment and obtained similar results leading to identical conclusions about price discovery.

the superscripts indicate whether the quote is a bid or an ask. We use Eqs. (5) and (6) only with short-term, near-the-money options on stocks for which ex-dividend dates of large dividends do not fall during the lives of the options. For these options, the early exercise premia are likely to be small, reducing the importance of any errors in our estimate of $v_t(K,T)$.

Specifically, we use only options that satisfy the following criteria:

(a) The remaining time to expiration T–t is between 10 and 70 calendar days, inclusive.
(b) The option is within 6% of being at-the-money, i.e., the option moneyness $\ln(S/K)$ satisfies $|\ln(S/K)| \leq 0.06$, where S is the average of all trade prices during the day.
(c) The present value of the dividends with ex-dividend dates during the remaining life of the option satisfies $PV_t(D\,(t,T)) < 0.05$.
(d) The bid price of the option is greater than or equal to 15 cents.

If more than one put-call option pairs satisfies the moneyness criterion (b), then all of them are considered independently, potentially resulting in the identification of more than one disagreement event in an underlying stock at a given time t. About 8.6% of the disagreement events have another event (for a different call-put pair based on the same underlying stock) within the next 10 seconds, and about 14% have another event for a different call-put pair within the next 2 minutes. These percentages would be higher but for the fact that often only one strike price satisfies the moneyness condition (b).

The first two criteria, (a) and (b), limit the value of the early exercise premia of both puts and calls, while (c) has the effect of including in the sample only the call options for which the call early exercise premium is small. Criterion (a) also greatly reduces the uncertainty about the dividend estimates used in (c), because dividends typically are announced about a month before the ex-dividend date and do not change from quarter to quarter. The criteria (a) and (b) also have the effect of eliminating many of the less liquid options from the sample, which is desirable.

Criterion (d) screens out many cases in which the option bid price becomes unavailable during the evaluation period. Such cases prevent us from computing the change in the option price. This criterion also contributes to screening out less liquid options.

For each underlying stock and ETF, we search for price disagreement events by comparing the option-implied stock bid and ask prices computed using Eqs. (5) and (6) to the actual bid and ask prices. In this process, we do not consider the quote updates from the first and last 5 minutes of each trading day. Skipping the first 5 minutes of the trading day avoids opening rotations and also allows the markets some time to aggregate information accumulated from the previous close. We cannot use disagreements that begin near the close of trading because we are interested in quote changes subsequent to the disagreements. For all other quote updates in the sample, a potential price disagreement event is identified at time t if there is a call-put pair that satisfies criteria (a)Å(d) and also

(i) S_t^{bid} – implied $\text{Ask}_t \geq \$0.02$ and (S_t^{bid} – implied Ask_t)/ $S_t^{bid} \geq 0.05\%$; or
(ii) Implied $\text{Bid}_t - S_t^{ask} \geq \$0.02$ and (implied $\text{Bid}_t - S_t^{ask}$)/ $S_t^{ask} \geq 0.05\%$.

These conditions imply that the bid-ask ranges of the actual and option-implied stock prices do not overlap, but rather are separated by a distance that is at least the greater of \$0.02% and 0.05% of the stock price. The absolute threshold \$0.02 is relevant for low-price stocks, while the relative threshold 0.05% is relevant for high-priced stocks.[6] In (i) the price is greater than the implied price, and we call such events {P > IP}-type disagreements; in (ii) the implied price exceeds the price, and we call such events {IP > P}-type disagreements. In the {P > IP}-type disagreements, the actual stock price should decrease or the option-implied price should increase to eliminate the disagreement, or both. In the {IP > P}-type disagreements, the stock price should increase or the implied price should decrease, or both.

We construct the final sample by applying several filters to the events that satisfy (i) and (ii). Any potential disagreement that is triggered within a 5-minute period following the start of any previous disagreement event is discarded. Events in which the bid-ask spread for the stock, call option, and put option are not all strictly positive are discarded, and all events with duration of less than half a second are discarded.[7] These filters ensure that the disagreement events that appear in the sample are not merely market locks, caused by differences in latency, or blinking limit orders that are displayed for only a few milliseconds, or otherwise so short-lived that investors are unable to trade on them. Finally, we do not search for disagreement events during the first and last 5 minutes of each trading day.

Our approach involves comparing the movement of both the actual and option-implied stock price quotes following the beginning of a disagreement event with the movement of the corresponding quotes following otherwise similar control events during which there is no disagreement about the underlying stock prices. In doing this we examine the quotes either 30 or 60 seconds after the beginning of the events, which we refer to as the evaluation period. Thus, we also associate with each disagreement event the stock and option-implied quotes prevailing 30 and 60 seconds subsequent to the beginning of the event.

Fig. 1 illustrates the quote dynamics during a disagreement event. Through the time 10:00:36.200, ABC's ask price is greater than or equal to \$60.00 and the option-implied quotes are \$60.00/\$60.10. The relative threshold

[6] We obtain similar results and identical conclusions if we use a threshold of the greater of 4 cents or 0.1% of the stock price.

[7] The number of these is about 10% of the sample size. Including them in the analysis strengthens the results.

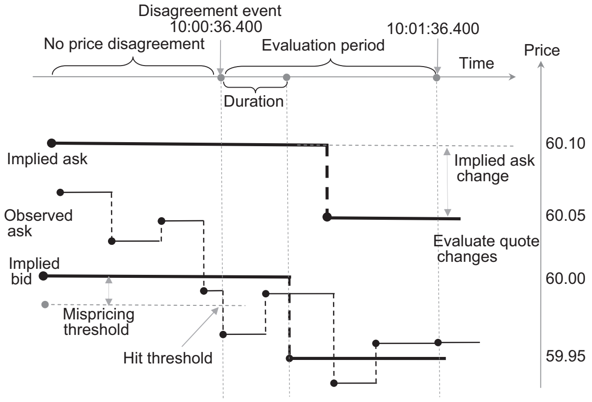

Fig. 1. Illustration of a price disagreement event. Through time 10:00:36.200, ABC's ask price is greater than or equal to $60.00 and the option-implied quotes are $60.00/$60.10. At the quotes prevailing through 10:00:36.200 the mispricing is only 1 cent and the disagreement event is not triggered. Then, two hundred milliseconds later, at 10:00:36.400, the actual ask quote changes from $59.99 to $59.97. The threshold of 3 cents is met and an {IP > P}-type price disagreement event is triggered. The quotes at 10:01:06.400 and 10:01:36.400 (30 and 60 seconds after the beginning of the disagreement event, respectively) are used in the analysis of quotes changes subsequent to the disagreement event.

is 3 cents (the absolute threshold is 2 cents), and it is the relevant one. At the quotes prevailing through 10:00:36.200 the mispricing is only 1 cent and the disagreement event is not triggered. Then, two hundred milliseconds later, at 10:00:36.400, the actual ask quote changes from $59.99 to $59.97. The threshold of 3 cents is met, an {IP > P}-type price disagreement event is triggered, and the time and quotes are stored. In addition, the quotes at 10:01:06.400 and 10:01:36.400 (30 and 60 seconds after the beginning of the disagreement event, respectively) are stored for the analysis of quotes changes subsequent to the disagreement event.

4.2. *Matched control sample*

We need to use a matched control sample for two reasons. First, during the evaluation period new information might arrive, causing prices to change. For example, does a stock price change during a disagreement event mean that the stock market has taken into account the option-implied stock prices, or is it simply that new information has arrived and been reflected in the stock price? A natural way to control for the arrival of new information and, thus, distinguish between these two explanations for the stock price movement is to compare the stock price changes during disagreement events with the stock price changes during a matched sample of otherwise similar market conditions for which no disagreement exists.

The same issue is relevant for the option market. Perhaps options quotes always move in a particular fashion, even absent price disagreement. For example, during disagreement events the option-implied bid-ask spread is usually smaller than its full sample average, and its reversion toward the full sample mean mechanically reduces the magnitude of the disagreement. Thus, at least part of the decrease in the disagreement is due to mean reversion in the bid-ask spread and not due to the options market following the stock market. Therefore, it is necessary to assess what are the normal changes in the option-implied quotes during the evaluation period. As in the previous case, a solution is to compare option dynamics under similar market conditions with and without disagreement. The question, then, is what variables should be used for matching?

Based on event mechanics, the option-implied bid-ask spread and pre-event stock returns are good candidates. Fig. 1 illustrates this. The actual stock bid-ask spread is narrow, consistent with the data. Thus, the stock price drifts inside a wide stripe of option-implied quotes, and a disagreement occurs when it gets outside the stripe. The disagreement is more likely to occur when the option-implied bid-ask spread is smaller than its average, because the stock has to move a smaller distance to get outside the option-implied quotes. Thus, the option-implied spread is an important variable to use in matching. As for additional variables, in our data most of the disagreement events are initiated by pre-event stock market returns. Because large stock price returns might by themselves predict subsequent movements in stock or option prices, we also use stock returns in both the 2 minutes and 10 seconds prior to the beginning of the event. These pre-event returns are also proxies for stock volatility at the time of the event.

To facilitate identification of the matching observations, the sample period is divided into periods of 40 trading days. For each disagreement event, a sample of potential matches from that underlying stock is constructed based on quote snapshots at 1-minute frequency from the 40-day period that contains the disagreement event. We exclude potential matches in which the stock, call, and put bid-ask spreads are not strictly positive, and we also exclude potential matches that occur within 5 minutes of the start of any of the disagreement events. Finally, we require that matches for the {P > IP}-type have an actual quote midpoint greater than the midpoint of the option-implied quotes and vice versa for the {IP > P}-type.

Each disagreement event is matched with exactly three observations from the set of potential matches. Selecting three matches for every disagreement event reduces sampling variation in the control sample. We use the Mahalabonis metric to select the matches.[8]

Panel A of Fig. 2 shows the behavior of normalized option-implied bid and ask prices and actual stock midpoint for both the disagreement and control samples during a four hundred second window surrounding the triggering of the {P > IP}-type disagreement events and the start of the corresponding control events, respectively.

[8] The Mahalanobis distance (Mahalanobis, 1936) between two vectors x and y with the same distribution is $d(x,y) = \sqrt{(x-y)'\Sigma^{-1}(x-y)}$, where $x,\ y \in \Re^N$ and $\Sigma \in \Re^{N\times N}$ is the covariance matrix. It is a scale-invariant dissimilarity measure that generalizes the Euclidian distance metric to include the covariance matrix Σ. When Σ is equal to the identity matrix I, the Mahalanobis distance is identical to the Euclidean distance.

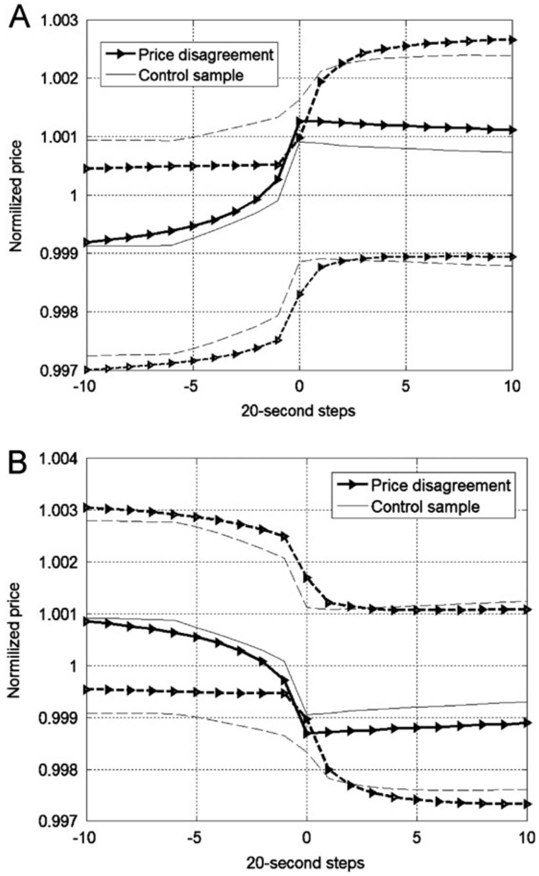

Fig. 2. Average levels of the normalized actual and option-implied stock prices around the disagreement events. The prices are normalized by dividing each price by the overall mean of the implied bid, implied ask, and stock midpoint observed at 20-second intervals during the four hundred-second window. Thick lines with arrows are for the price disagreement sample, thin lines are for the control sample, dashed lines are the option-implied bid and ask prices, and solid lines are the actual stock bid-ask midpoints.

The prices are normalized by dividing each price by the overall mean of the implied bid, implied ask, and stock midpoint observed at 20-second intervals during the window, so that a normalized price of 1 means that the price is equal to the overall mean price. The thick lines with arrows are for the price disagreement sample, the thin lines are for the control sample, dashed lines are the option-implied bid and ask prices, and solid lines are the actual stock bid-ask midpoints. The actual stock price for the disagreement events (the thick solid line with arrows) rises very rapidly before time 0, the beginning of the disagreement event, and crosses above the dashed solid line (the option-implied ask) shortly before the disagreement event. At time 0 it equals or exceeds the option-implied ask plus the threshold, triggering the disagreement event. The figure shows that the disagreement is typically triggered by the rapid increase in the actual stock price relative to the option implied ask, and not by a decline in the option-implied ask. The disagreement is then closed shortly after time 0 when the option-implied ask catches up with the stock price and returns to its typical position above the actual stock price.

The behavior of the stock price in the control sample is similar to the behavior in the disagreement sample, because matching on both 2-minute and 10-second pre-event returns captures the convexity of the stock price trajectory in the pre-event period. The main difference between the control and disagreement samples is that the control sample option-implied bid and ask prices increase more quickly, so a disagreement is not triggered.

Panel B shows that the average dynamics during the $\{IP > P\}$-type events are almost a mirror of the $\{P > IP\}$-type events. The disagreement events are triggered by stock price movements in which the stock price, the thick solid line with arrows, declines more rapidly than the implied bid and ask. Subsequent to the triggering of the event, the option market catches up.

4.3. Broader sample and matched control sample for the broader sample

To address the potential concern that the results based on the primary sample are due to anomalous option market behavior during violations of the law of one price, we also construct and use a much larger sample based on a broader definition of a disagreement. For this sample we do not use the filters described above, but instead include every option for which the bid-ask spread was strictly positive at the time the disagreement is triggered and the absolute value of delta (as computed by OptionMetrics) as of the close of trading on the previous day was between 0.1 and 0.9, inclusive. By including options relatively far away from the money, this broader sample also addresses the potential concern that we fail to find meaningful price discovery in the primary sample because price discovery occurs in out-of-the-money options and the primary sample includes only near-the-money options.

For the broader sample, a disagreement occurs when the model value of an option computed using the current stock price and a lagged estimate of the implied volatility is at least one penny outside the option's bid-ask range. Specifically, as a preliminary step for each option every 2 minutes we compute the Black-Scholes-Merton implied volatility using the current option and stock bid-ask midpoints. Given this set of implied volatilities for each option at 2-minute frequency, every time one of the stock or option bid or ask quotes changes we use the current stock bid-ask midpoint and the average of the implied volatilities for the preceding 15 2-minute intervals to compute the Black-Scholes-Merton value of the option.[9] If the computed model value is outside the quoted bid-ask range by at least one penny, we trigger a disagreement event. This broader

[9] This exploits the slow-moving nature of implied volatilities; i.e., there is little high-frequency movement in implied volatilities. Because we ignore the American feature in both computing the implied volatilities and in using them to compute the model values, the errors from ignoring the early exercise premia at the two steps offset each other. Even using the Black-Scholes-Merton formula, the computation of the implied volatilities at high frequency is burdensome.

sample includes almost 55 times as many disagreement events as the primary sample.

To construct the control sample for the broader sample, we match on the option bid-ask spread, the 2-minute pre-event stock return, the square root of the days remaining to expiration, and the absolute value of the option delta, and we use the Mahalabonis metric. We exclude potential matches that fall within a five-minute window surrounding the beginning of the disagreement event. For more than 85% of the events in the broader sample, the matched option is the same option as the one that triggered the disagreement event, in 50% of the events the matched observation is from the same trading day, and in 48% of the cases the matched observation is both the same option and from the same trading day.

4.4. Summary statistics

Table 1 presents some summary information about the primary sample disagreement events, on a stock-by-stock basis. The first column shows the stock ticker symbol. The second through sixth columns are for the $\{P>IP\}$-type disagreement events, and the seventh through 11th are for the $\{IP>P\}$-type events. The last row shows either the sum or mean of the stock-by-stock quantities, as appropriate.

There are 38,979 events of the $\{P>IP\}$-type and 42,045 of the $\{IP>P\}$-type in the primary sample, a total of 81,024 events. The largest event suppliers are eBay and Amazon, with more than eight thousand events each. Some of the stocks (e.g., America Online and Calpine) dropped from the sample before the end of the sample period. Thus, there are few events for them. The broader sample, for which summary statistics are not tabulated, contains 2,315,862 $\{P>IP\}$-type events and 2,132,490 $\{IP>P\}$-type events, a total of 4,448,352 events. eBay and Qualcomm (QCOM) are the leading events suppliers in the broader sample, and Amazon is the third leading supplier.

The primary sample disagreements occur on average more than twice per stock, per day. The disagreement frequency falls from about four per day in the beginning of the sample period to less than one per day at the end, with an overall average of more than two per stock per day. There are two possible drivers for this trend. First, volatility as measured by the Chicago Board Options Exchange (CBOE) Market Volatility Index (VIX) fell from 25% in April 2003 to 12% in October 2006. Next, improvements in technology could have made it easier for option market makers to avoid price disagreements. Based on the broader definition of a disagreement event that results in a sample of about 4.5 million events, disagreements occur on average more than 130 times per stock, per day. These disagreement events are important for the option market. Using the broader definition of a disagreement event, 4.5% is a lower bound on the fraction of total (across all options) option trading volume that comes from trades in the option suffering a disagreement that occur within 30 seconds of the triggering of the disagreement event. Using a 60-seconds window, 6.7% is a lower bound on the fraction of total option trading volume that occurs in the disagreement options. These percentages are lower bounds on the fractions of option trading volume that occur during the disagreement events because we filter out subsequent events that begin within 5 minutes following the triggering of any event. Thus, option trading volume during the subsequent events in not counted as disagreement volume but is included in the total option volume that appears in the denominators of these calculations.

The third and eighth columns show the median event durations (in seconds), by stock. The means (across stocks) of the median event durations are 16.4 and 18.3 seconds for the $\{P>IP\}$ and $\{IP>P\}$-type disagreement, respectively, with the smallest and largest median durations ranging from 4.2 seconds ($\{IP>P\}$-type events for QCOM) to 73 seconds ($\{IP>P\}$-type events for CPN). Disagreement events have a short life, with 70% of disagreements disappearing in the first 20 seconds and 90% being eliminated within 70 seconds. Because we exclude from the sample disagreements with durations of less than one-half second and some (and perhaps most) of these short-lived disagreements are valid disagreements and not due to differences in latency or data reporting problems, the table somewhat overstates the typical lengths of disagreement events. Our choice of evaluation periods of 30 and 60 seconds is motivated by this information about the average lives of the events.

The columns headed "Implied spread" shows the average spread between the option-implied bid and ask prices. The overall average implied spread is 11 cents, with many option-implied quotes displaying an implied spread of 10 cents, the minimum possible implied spread. These wide average implied spreads justify thinking intuitively of the stock price drifting inside the stripe of option-implied quotes.

For both event types, both the pre-event stock returns are substantially different from zero. Specifically, the 2-minute pre-event returns are 0.22% for the $\{P>IP\}$-type and -0.22% for the $\{IP>P\}$-type, while the ten-second pre-event returns are 0.12% and -0.11%, respectively. These pre-event returns are consistent with the fact that most price disagreements are precipitated by stock price changes. In fact, 88% of the disagreements are initiated by the stock in the sense that a change in a stock price quote was the immediate cause of the disagreement, and only 12% of the disagreements are initiated by the options. The pre-event returns in the broader sample are also positive and negative for the $\{P>IP\}$-type and $\{IP>P\}$-type disagreements, respectively, though of smaller magnitude than the pre-event returns for the primary sample events.

5. Hypotheses and main results

This section presents the main results on the differences between the distributions of the quote changes in the disagreement and control samples, and explains how the differences bear on our hypotheses about price discovery in the stock and option markets.

5.1. Distributions of quote changes

We begin with a visual examination of the differences between the distributions of the quote changes in the disagreement and control samples. For the $\{P>IP\}$-type disagreements, the option implied ask quote is less than the actual bid. The hypothesis that the options follow the stock implies that the distribution of the change in the option-implied ask quote for the disagreement sample is to the right of the distribution for the control sample, while the hypothesis that the options do not follow the stock implies that the distribution of the change in the option-implied ask quote for the disagreement sample is either identical to or to the left of the distribution for the control sample. The hypothesis that the stock follows the options implies that for these {P > IP}-type disagreements the distribution of the change in the actual bid quote for the disagreement sample is to the left of the distribution for the control sample, while the hypothesis that the stock does not follow the options implies that the distribution of the change in the actual bid quote for the disagreement sample is either identical to or the right of the distribution for the control sample.

For the {IP > P}-type disagreements, the option-implied ask quote is greater than the actual bid and the predictions for the differences in the distributions are opposite those above. Specifically, the hypothesis that the options follow the stock implies that the distribution of the change in the option-implied ask quote for the disagreement sample is to the left of the distribution for the control sample, and the hypothesis that the stock follows the options implies that the distribution of the change in the actual bid quote for the disagreement sample is to the right of the distribution for the control sample.

Fig. 3 consists of six panels comparing the quote change distributions over an evaluation period of 1 minute for the treatment and control samples aggregated across all stocks. In constructing the graphics, for simplicity we replace the actual stock bid price with the bid-ask midpoint because the bid–ask spread in the stock market was typically only 1 cent and the distribution of changes in the actual bid-ask midpoint cannot be visually distinguished from the distributions of changes in the actual bid and ask prices. The top three panels (A–C) are for the {P > IP}-type disagreements and compare the distributions of changes in the option-implied bid, option-implied ask, and actual stock midpoint for the disagreement and control samples. In each panel, the distribution of quote changes in the disagreement sample is represented by the light-shaded bars, and the distribution of quote changes in the control sample is represented by the dark-shaded bars. The bottom three panels (D–F) also compare the distributions of changes in the option-implied bid, option-implied ask, and actual stock midpoint quotes between the disagreement and control samples, but for the {IP > P}-type disagreements. In all cases quote changes are measured in US dollars, e.g., 0.2 is 20 cents. The distributions of changes in the option-implied bid and ask prices are discrete, as the tick size is 5 cents.

For the {P > IP}-type disagreements (Panels A–C) the most noticeable difference in the distributions is in Panel B, which compares the distributions of the changes in the implied ask prices for the treatment and control samples. For the control sample (dark-shaded bars), the implied ask stays the same or decreases more than 60% of the time; it increases and moves toward the actual bid in fewer than 40% of the events. The finding that the implied ask moves toward the bid in some of the events should not be surprising, as in both the disagreement and control samples the option-implied bid-ask spread is less than its

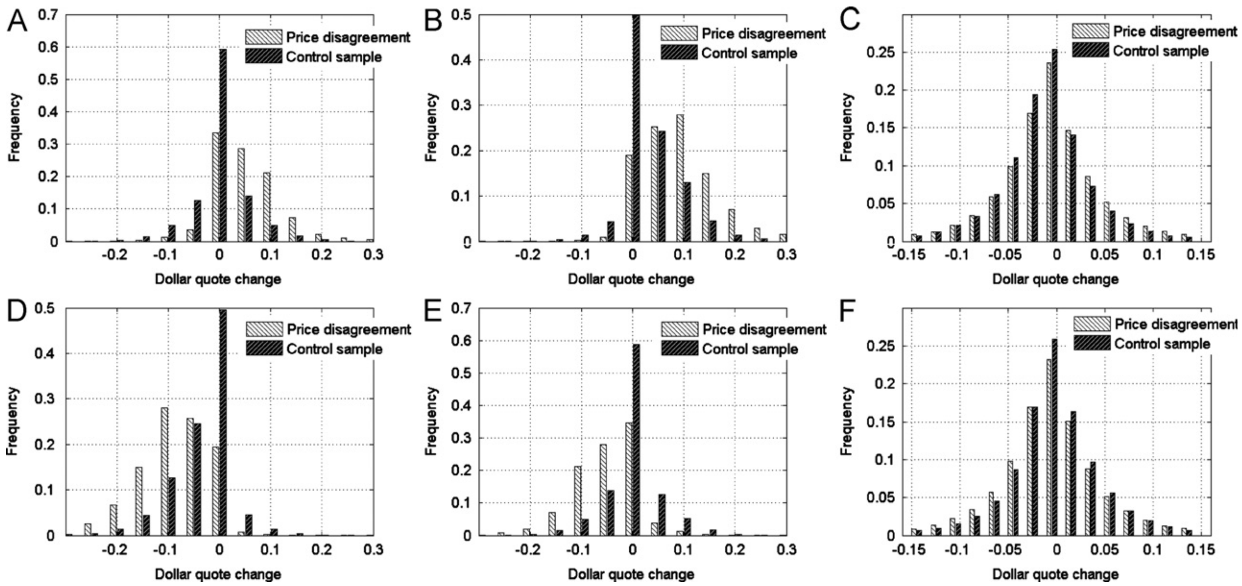

Fig. 3. Comparisons of the distributions of quote changes for the primary disagreement and control samples. Panels A–C compare the distributions of changes in the option-implied bid, option-implied ask, and actual stock midpoint, respectively, for the disagreement and control samples for the {P > IP}-type disagreements, using an evaluation period of 1 minute. In each panel, the distribution of quote changes in the disagreement sample is represented by the light-shaded bars, and the distribution of quote changes in the control sample is represented by the dark-shaded bars. Panels D–F compare the distributions of changes in the option-implied bid, option-implied ask, and actual stock midpoint quotes, respectively, between the disagreement and control samples for the {IP > P}-type disagreements. In all cases, quote changes are measured in US dollars. Panels A, B, D, and E use a bin size of 5 cents, which is the tick size in the options market. Panels C and F showing the distribution of changes in the stock midpoint use a bin size of 2 cents.

unconditional average value, so some spread widening is to be expected. By construction in the control sample, the observed midpoint is closer to the implied ask than to the bid. In contrast, the distribution of changes in the treatment sample shows a much higher frequency of increases in the implied ask price. The implied ask stays the same or decreases in about 20% of the treatment observations and increases in more than 80%. The mode of the distribution of the treatment sample is 10 cents, and the mean is close to 8 cents. Thus, the implied ask typically changes by more than enough to eliminate the initial pricing disagreement. Importantly, the whole treatment distribution is to the right of the control distribution, implying that, for the disagreement events, the change in the implied ask price is greater at every percentile.

Panel A shows that it is also the case that the distribution of changes in the option-implied bid prices for the disagreement sample is shifted to the right relative to the distribution for the control sample. For the control sample, the probability that the option-implied bid price does not change is about 60%, and the distribution is approximately symmetric. This is less movement than was found in the control sample ask prices, which is unsurprising because in the control sample by construction the stock midpoint is closer to the implied ask than to the bid. In contrast, the probability that the option-implied bid increases during the evaluation period is about 60%. Combined with the movement of the implied ask, this shift in the distribution of the bid prices provides strong evidence that the options market changes its assessment of a fair price during the disagreement events. This implies that the stock market participates in price discovery.

Examining the distributions of changes in the stock price allows us to assess whether the options market participates in price discovery. If the stock price follows the options, then for the disagreement sample the distribution of changes in the actual stock quote midpoint should be shifted to the left relative to the control sample. Panel C reveals that there is no such shift. Instead, the distributions for the treatment and controls samples are similar. If anything, the treatment distribution is shifted slightly to the right of the distribution for the control sample and displays slightly more dispersion. This is evidence that the stock market does not follow the options, but rather is not affected by the disagreement.

The distributions for the {IP > P}-type disagreements in which the option-implied bid exceeds the actual ask are shown in Panels D–F and are almost a mirror image of the figures for the {P > IP}–type in Panels A–C. The distributions of changes in the option-implied bid and ask quotes for the price disagreement samples are shifted to the left relative to the corresponding distributions for the control samples, providing evidence that the options follow the stock. The distributions of changes in the actual stock prices are very similar for the treatment and control samples, indicating that the stock does not follow the options. If anything, the treatment distribution is shifted slightly to the left of the control distribution, also mirroring the results for the {P > IP}-type disagreements.

The distributions of quotes changes in these six panels are consistent with the hypotheses that option price quotes follow stock price quotes, but not vice versa. For both disagreement types, both option-implied bid and ask prices tend to change in the directions that reduce or eliminate the disagreement. In striking contrast, stock prices do not change in directions that reduce or eliminate the disagreement. To the extent that the distributions changes in stock prices differ between the treatment and control samples, the stock prices change in the directions that increase the magnitude of the disagreement.

5.2. Comparison of mean differences in the primary sample

For the primary sample {P > IP}-type disagreements, the option-implied ask quote is less than the actual bid. This suggests a focus on movements in the option-implied ask and the actual bid because these are the quotes most likely to change to eliminate the disagreement. However, for simplicity we use the stock midpoint instead of the bid price. This has no effect on our analysis as the bid-ask spread is typically only 1 cent. The mean changes in the option-implied ask quotes and actual stock midpoints are estimated using

$$\text{Implied Ask}_{t+\tau} - \text{Implied Ask}_t = a_0 + a_1 D_t + a_2 X_t + \varepsilon_{t+\tau}, \tag{7}$$

and

$$\text{Stock Midpoint}_{t+\tau} - \text{Stock Midpoint}_t = b_0 + b_1 D_t + b_2 X_t + \varepsilon_{t+\tau}, \tag{8}$$

where τ is the length of the evaluation period, D is a disagreement dummy that takes the value one for disagreement events, and X is a vector of control variables relevant for the option-implied quotes. The estimates of the coefficients a_1 and b_1 are estimates of the mean differences between the treatment and control sample; i.e., they are estimates of the effect of the disagreement on subsequent mean changes in quotes. If the stock market quotes contain information not yet reflected in the option quotes, then $a_1 > 0$, interpreted as the options follows the stock; otherwise, $a_1 \leq 0$. If the options market quotes contain information not yet reflected in the stock quotes, then $b_1 < 0$, interpreted as the stock follows the options; otherwise, we expect $b_1 \geq 0$.

For the primary sample {IP > P}-type disagreements, we focus on the option-implied bid and the actual stock midpoint. The mean changes in the option-implied bid and actual stock midpoint are estimated using

$$\text{Implied Bid}_{t+\tau} - \text{Implied Bid}_t = a_0 + a_1 D_t + a_2 X_t + \varepsilon_{t+\tau} \tag{9}$$

and Eq. (8), where $a_1 < 0$ means that the options follow the stock and in Eq. (8) $b_1 > 0$ now means that the stock follows the options.

If the regressions do not include the vector of control variables X, then these regressions and hypotheses are equivalent to a classic average treatment effect analysis. Table 2 presents summary results for changes in the implied bid, implied ask, and stock midpoint based on evaluation periods of both 30 seconds and 1 minute.[10] (The quote changes are calculated at the end of the evaluation periods.) Specifically, the table shows the across-stock means and medians of the stock-by-stock

[10] The Internet Appendix reports detailed stock-by-stock results for the 1-minute evaluation period.

Table 2

Comparison of mean quote changes in the primary disagreement and control samples for 30-second and 1-minute evaluation periods.

The table presents summary information about the mean changes (in cents) of the option-implied bid quote, option-implied ask quote, and actual stock midpoint for the disagreement and the control samples, for both the {P > IP}-type and {IP > P}-type disagreements. The columns headed "Mean" and "Median" present the mean and median, respectively, across the 39 stocks and exchange-traded funds (ETFs), of the mean quote changes computed for the 39 stocks and ETFs. The columns headed "4th" and "36th" present the 4th and 36th of the mean quote changes across the 39 stocks and ETFs. The number of disagreements for each stock is reported in Table 1. Because each disagreement event is matched to three control events, the control samples are three times the sizes of the treatment samples.

Price change and sample	{P > IP}-type disagreements				{IP > P}-type disagreements			
	Mean	Median	4th	36th	Mean	Median	4th	36th
Option-implied bid								
30-second evaluation period								
Treatment sample	4.2	4.1	3.0	5.5	−6.3	−5.7	−10.4	−3.6
Control sample	0.4	0.4	0.1	0.5	−2.3	−2.1	−3.9	−1.1
1 minute evaluation period								
Treatment sample	4.8	4.6	4.1	6.1	−7.7	−6.8	−11.7	−5.3
Control sample	0.3	0.4	−0.1	0.7	−2.7	−2.6	−4.1	−1.5
Option-implied ask								
30-second evaluation period								
Treatment	6.6	5.8	4.0	9.5	−4.0	−4.1	−5.3	−2.9
Control	2.3	1.9	0.9	3.8	−0.3	−0.2	−0.5	0.0
1-Minute evaluation period								
Treatment	8.1	7.1	5.3	11.4	−4.6	−4.6	−5.9	−3.3
Control	2.7	2.6	1.2	4.2	−0.2	−0.2	−0.7	0.2
Stock midpoint								
30-second evaluation period								
Treatment	0.1	0.1	−0.3	0.4	−0.1	−0.2	−0.5	0.3
Control	−0.1	0.0	−0.4	0.2	0.1	0.0	−0.2	0.4
1 minute evaluation period								
Treatment	0.1	0.1	−0.3	0.4	0.0	−0.2	−0.6	0.4
Control	−0.2	−0.2	−0.6	0.2	0.2	0.2	−0.2	0.7

average quote changes and also presents the 4th and 36th average quote changes, ranking the averages from smallest to largest. The left-hand side of the table is for the {P > IP}-type disagreements, and the right-hand side is for the {IP > P}-type disagreements.

Examining the results for the {P > IP}-type disagreements, in the treatment sample for the 1-minute evaluation period, the across-stock mean and median of the stock-by-stock average changes in the implied bid are 4.8 and 4.6 cents, respectively, and the 4th and 36th values are 4.1 and 6.1 cents, respectively. Although it is not shown in the table, the minimum of the average changes across stocks is 2.4 cents. For the control sample, the across-stock mean and median of the average changes in the implied bids are only 0.3 and 0.4 cents, respectively, and the 4th and 36th values are −0.1 and 0.7 cents, respectively.

The across-stock mean and median changes in the treatment sample option-implied ask quotes are somewhat larger, being 8.1 and 7.1 cents, respectively, for the one-minute evaluation period. For the control sample, the corresponding across-stock mean change in the implied ask is 2.7 cents, reflecting some spread widening because disagreements are more likely to occur when the option-implied bid-ask spread is smaller than average. The average difference in the implied ask quote between the treatment and control samples is 8.1–2.7 = 5.4 cents, larger than the initial mispricing.

If the stock followed the options, the average changes in the stock midpoints in the treatment sample would be negative and less than the corresponding average changes for the control sample. Inconsistent with this hypothesis, the across-stock mean and median average changes in the stock midpoint are both 0.1 cents, while in the control sample, the across-stock mean and median changes are both −0.2 cents. The 4th and 36th largest treatment sample average changes also exceed the corresponding average changes in the control sample. In fact, while not shown in the table, for every stock the average change in the stock midpoint in the treatment sample exceeds the average change in the control sample. Together with the results for the option-implied quotes, these results imply that for the {P > IP}-type disagreements the options follow the stock and the stock does not follow the options.

The right-hand half of the table presents the corresponding average quote changes for the {IP > P}-type disagreements. The across-stock means and medians of the stock-by-stock average changes in the option-implied bid and ask are now negative and less than the corresponding control sample across-stock means and medians, implying that in this case also the options follow the stocks. (For the {IP > P}-type disagreements, the option-implied stock prices exceed the actual stock prices, so that if the options follow the stock the average changes in the option-implied bid and ask prices will be negative.) The across-stock mean change in the implied bid is −7.7 cents, close to the magnitude of the across-stock mean change in the implied ask of 8.1 cents for the {P > IP}-type disagreements, and the across-stock mean change in the implied bid is −4.6 cents, close to the magnitude of the across-stock mean change in the implied bid of 4.8 cents for the {P > IP}-type disagreements. The across-stock mean change in stock prices is zero and less than the across-stock change in the control sample, inconsistent with the hypothesis that the stocks follow the options. Overall, this table produces strong evidence that if prices disagree, the options market adjusts to eliminate mispricing and the stock market does not adjust. It seems worth emphasizing that the average changes in the option prices are large, and typically larger than the extent of the disagreement, while the average change in stock prices are close to zero. To the extent that the average changes in stock prices are nonzero, the stock prices move to widen the disagreement rather than reduce it.

The results for the 30-second evaluation period show that our findings are not sensitive to the length of the evaluation period. Consistent with the previous results based on a 1-minute evaluation period, the treatment sample changes in the option-implied quotes are positive and greater than the corresponding changes in the control sample for the {P > IP}-type disagreements, while the treatment sample changes in the option-implied quotes are negative and smaller than the corresponding changes in the control sample for the {IP > P}-type disagreements. Unsurprisingly, the magnitudes of the average changes are smaller for the shorter 30-second evaluation period. Also consistent with the previous results, the average changes in the actual stock prices are close to zero.

5.3. Conditional mean differences in the primary sample

The high-frequency data provide control observations that match very closely the characteristics of the treatment observations. Nonetheless, the matching is not perfect. In such circumstances, Imbens and Wooldridge (2009) suggest controlling for the differences between the treatment and control samples by estimating regressions that include as covariates the variables used in matching the samples. Following this advice, we use a vector of control variables X consisting of the three variables used in matching (the option-implied bid-ask spread at the beginning of the event and the 2-minute and 10-second pre-event returns on the underlying stock) and an additional variable consisting of the signed order imbalance in the stock during the 2 minutes prior to the beginning of the event. We estimate the Regressions (7)–(9) separately for each underlying stock and ETF.

Table 3 presents some summary information about the 39 estimates of the dummy variable for the implied bid, implied ask, and stock midpoint based on both the 30-second and 1-minute evaluation periods. It shows the means and medians across stocks of the stock-by-stock estimates of the dummy variable, and also presents the 4th and 36th estimates, as well as the associated t-statistics. In the column headed "Mean," the t-statistic is the mean of the t-statistics for the 39 stocks and ETFs. The other columns report the t-statistic of the median, 4th, and 36th estimates. Again the left-hand side of the table is for the {P > IP}-type disagreements, and the right-hand side is for the {IP > P}-type disagreements. For the {P > IP} type and using the one-minute evaluation period, for the implied ask the mean across stocks of the estimates is 5.2 cents, and the estimates varies from 3.7 cents for the 4th to 7.2 cents for the 36th. The mean of the dummy coefficient estimates for the implied bid is 4.5 cents, only slightly smaller than the coefficient estimate in the regression

Table 3
Estimates of the mean effect of disagreement in the primary sample, controlling for the three variables used in matching and the stock market order imbalance.

The table presents summary information about the coefficient estimates on the disagreement dummy in regressions of quote changes on a constant, the disagreement dummy, and the three variables used in matching (the option-implied spread and the two-minute and ten-second pre-event returns), and a fourth control variable equal to the order imbalance in the stock during the two-minute pre-event period. The columns headed "Mean" and "Median" present the mean and median, respectively, across the 39 stocks and exchange traded funds (ETFs), of the coefficient estimates on the disagreement dummy and their associated t-statistics (in parentheses). The columns headed "4th" and "36th" present the 4th and 36th the coefficient estimates, again with the associated t-statistics. In the column headed "Mean," the t-statistic is the mean of the t-statistics for the 39 stocks and ETFs. The other columns report the t-statistic of the median, 4th, and 36th estimates, respectively. The t-statistics are based on White heteroskedasticity-consistent standard errors. The number of disagreements for each stock is reported in Table 1. Because each disagreement event is matched to three control events, sample sizes for individual regressions are four times the number of disagreements for each stock reported in Table 1.

Price change and evaluation period	{P > IP}-type disagreements				{IP > P}-type disagreements			
	Mean	Median	4th	36th	Mean	Median	4th	36th
Implied bid								
30 seconds	3.9	3.6	2.7	5.2	−4.0	−3.6	−5.9	−2.3
	(21.5)	(7.5)	(19.4)	(29.6)	(−20.1)	(−17.2)	(−12.9)	(−8.9)
1 minute	4.5	4.3	3.5	5.9	−4.9	−4.4	−7.2	−3.1
	(20.4)	(33.5)	(13.8)	(19.0)	(−20.9)	(−31.8)	(−23.4)	(−4.1)
Implied ask								
30 seconds	4.2	3.7	2.7	6.2	−3.8	−3.8	−5.2	−2.6
	(19.1)	(7.1)	(3.5)	(16.2)	(−22.2)	(−22.0)	(−52.5)	(−4.5)
1 minute	5.2	4.4	3.7	7.2	−4.5	−4.3	−5.9	−3.4
	(20.5)	(31.6)	(29.2)	(14.6)	(−20.9)	(−35.0)	(−30.7)	(−6.4)
Stock midpoint								
30 seconds	0.2	0.1	−0.3	0.7	−0.2	−0.3	−0.7	0.3
	(2.0)	(0.5)	(−1.3)	(7.1)	(−2.2)	(−2.7)	(−6.8)	(2.1)
1 minute	0.2	0.3	−0.3	0.7	−0.2	−0.3	−0.7	0.2
	(1.8)	(1.6)	(−0.9)	(3.8)	(−2.0)	(−2.1)	(−5.5)	(0.5)

explaining the ask quote. Thus, an option market marker moves both bid and ask by similar amounts, conditional on matching variables. These estimates are highly significantly different from zero at conventional levels, though one should recognize that the sample consists of 324,096 observations.

For the actual stock midpoint, the mean of the coefficient estimates is 0.2 cents, with an average t-statistic of only 1.8 despite the large sample size. Importantly, to the extent that the movement of the stock price is nonzero, this provides evidence that the actual stock price moves in the direction of increasing the disagreement, not decreasing it. For the {IP > P}-type, the results are again the mirror image of those for the {P > IP}-type disagreements.

5.4. Mean and conditional mean differences in the broader sample

Tables 4 and 5 present the results for mean differences and conditional mean differences in the broader sample of about 4.5 million disagreement events. An important difference between the results in Tables 4 and 5 and the corresponding results for the primary sample in Tables 2 and 3 is that in the broader sample we do not use the option quotes to compute option-implied bid and ask prices for the stock, but rather work with the actual option bid and ask prices. Thus, in the analysis of the broader sample the regression models (7) and (9) for the option-implied ask and bid (of the stock) are replaced by regression models for the option ask and bid,

$$\text{Option Ask}_{t+\tau} - \text{Option Ask}_t = a_0 + a_1 D_t + a_2 X_t + \varepsilon_{t+\tau} \quad (10)$$

and

$$\text{Option Bid}_{t+\tau} - \text{Option Bid}_t = a_0 + a_1 D_t + a_2 X_t + \varepsilon_{t+\tau}. \quad (11)$$

Table 4 shows the means and medians across stocks of the stock-by-stock average changes in the option bid, option ask, and stock midpoint and also presents the 4th and 36th average changes, for both 30-second and 1-minute evaluation periods. Again the left-hand side of the table is for the {P > IP}-type disagreements, and the right-hand side is for the {IP > P}-type disagreements. For the {P > IP}-type disagreements, increases in the option ask price resolve the disagreements, and the mean changes in the ask are 3.3 and 3.8 cents for the 30-second and 1-minute evaluation periods, respectively. These changes in the ask quotes are much larger than the corresponding changes in the control sample, which reflect some spread widening. Changes in the implied bid are smaller but still positive, in contrast to the (small) negative changes in the control sample that reflect spread widening. For the {IP > P}-type disagreements decreases in the option bid price resolve the disagreements, and the mean changes in the ask are −3.5 and −4.0 cents for the 30-second and 1-minute evaluation periods, respectively, which are larger than the corresponding changes in the control sample. Changes in the option ask are smaller, but negative, and contrast with the positive changes in the control sample. Also consistent with the previous results, the changes in the actual stock prices are close to zero and

Table 4

Comparison of mean quote changes in the broader disagreement and control samples for 30-second and 1 minute evaluation periods.

The table presents summary information about the mean changes (in cents) of the option bid quote, option ask quote, and actual stock midpoint for the disagreement and the control samples, for both the {P > IP}-type and {IP > P}-type disagreements. The columns headed "Mean" and "Median" present the mean and median, respectively, across the 39 stocks and ETFs, of the mean quote changes computed for the 39 stocks and exchange- traded funds (ETFs). The columns headed "4th" and "36th" present the 4th and 36th mean quote changes across the 39 stocks and ETFs. The number of {P > IP}-type disagreements per underlying stock range from 2551 to 199,122, and the number of {IP > P}-type disagreements per underlying stock range from 2551 to 195,500. Because each disagreement event is matched to three control events, the control samples are three times the sizes of the treatment samples.

Price change and sample	{P > IP}-type disagreements				{IP > P}-type disagreements			
	Mean	Median	4th	36th	Mean	Median	4th	36th
Option bid								
30-second evaluation period								
Treatment	1.1	1.1	0.7	1.7	−3.5	−3.4	−5.1	−2.1
Control	−0.5	−0.4	−0.9	−0.2	−0.6	−0.6	−1.1	−0.2
1 minute evaluation period								
Treatment	1.3	1.3	0.9	1.9	−4.0	−4.0	−5.5	−2.6
Control	−0.7	−0.6	−1.2	−0.3	−0.8	−0.7	−1.4	−0.3
Option ask								
30-second evaluation period								
Treatment	3.3	3.2	1.9	4.7	−1.1	−1.2	−1.8	−0.6
Control	0.6	0.5	0.2	1.0	0.5	0.4	0.2	0.9
1-minute evaluation period								
Treatment	3.8	3.7	2.4	5.3	−1.4	−1.4	−2.1	−0.8
Control	0.8	0.7	0.3	1.3	0.7	0.6	0.3	1.2
Stock midpoint								
30-second evaluation period								
Treatment	0.0	0.0	−0.3	0.3	0.0	0.0	−0.3	0.3
Control	−0.1	−0.1	−0.1	0.0	0.0	0.0	0.0	0.1
1 minute evaluation period								
Treatment	−0.1	−0.1	−0.4	0.2	0.1	0.0	−0.2	0.4
Control	−0.1	−0.1	−0.3	0.0	0.1	0.1	0.0	0.2

Table 5
Estimates of the mean effect of disagreement in the broader sample, controlling for the variables used in matching.

The table presents summary information about the coefficient estimates on the disagreement dummy in regressions of quote changes on a constant, the disagreement dummy, and the four variables used in matching: the option bid-ask spread, the 2-minute pre-event stock return, the square root of the days remaining to option expiration, and the option delta. The columns headed "Mean" and "Median" present the mean and median, respectively, across the 39 stocks and exchange-traded funds (ETFs), of the coefficient estimates on the disagreement dummy and their associated *t*-statistics (in parentheses). The columns headed "4th" and "36th" present the fourth and 36th coefficient estimates (ranked from smallest to largest), again with the associated *t*-statistics. In the column headed "Mean," the *t*-statistic is the mean of the *t*-statistics for the 39 stocks and ETFs. The other columns report the *t*-statistic of the median, 4th, and 36th estimates, respectively. The *t*-statistics are based on White heteroskedasticity-consistent standard errors. The number of {P > IP}-type disagreements per underlying stock range from 2551 to 199,122, and the number of {IP > P}-type disagreements per underlying stock range from 2204 to 195,500. Because each disagreement event is matched to three control events, the sample sizes for individual regressions are four times the number of disagreements for each stock.

Price change and evaluation period	{P > IP}-type disagreements				{IP > P}-type disagreements			
	Mean	Median	4th	36th	Mean	Median	4th	36th
Option bid								
30 seconds	1.6	1.5	0.9	2.5	−2.9	−2.9	−4.1	−1.7
	(115.3)	(109.2)	(52.8)	(231.2)	(−142.1)	(−139.7)	(−296.7)	(−80.1)
1 minute	2.0	1.9	1.3	3.0	−3.1	−3.2	−4.2	−2.2
	(116.5)	(113.7)	(59.7)	(196.6)	(−136.0)	(−145.4)	(−226.8)	(−71.9)
Option ask								
30 seconds	2.7	2.6	1.6	3.8	−1.6	−1.6	−2.7	−0.9
	(142.6)	(139.6)	(23.1)	(251.9)	(−108.8)	(−118.0)	(−191.8)	(−18.5)
1 minute	3.0	3.0	2.0	3.9	−2.1	−2.1	−3.2	−1.2
	(138.1)	(144.4)	(26.3)	(237.1)	(−116.3)	(−161.2)	(−231.6)	(−21.9)
Stock midpoint								
30 seconds	0.0	0.0	−0.2	0.4	0.0	0.0	−0.4	0.2
	(1.6)	(−1.4)	(−13.9)	(24.1)	(−1.1)	(−2.4)	(−28.8)	(12.7)
1 minute	0.0	0.0	−0.2	0.4	0.0	0.0	−0.4	0.2
	(0.7)	(−1.6)	(−14.2)	(16.8)	(−0.5)	(−1.7)	(−24.7)	(13.0)

are not meaningfully different from the changes in the control sample.

Table 5 presents summary information about the 39 estimates of the dummy variable in regressions with control variables that explain the option bid, option ask, and stock midpoint, for both the 30-second and one-minute evaluation periods. The control variables are the four variables used in matching, these being option bid-ask spread, the 2-minute pre-event stock return, the square root of the days remaining to option expiration, and the option delta. For the {P > IP}-type events, the inclusion of the control variables somewhat increases (decreases) the magnitudes of the estimated changes in the option bid (ask), and for the {P > IP}-type events it somewhat decreases (increases) the magnitudes of the estimated changes in the option bid (ask). However, the conclusion that the disagreements lead to economically significant changes in the option quotes is not affected. Also consistent with previous results, the mean and median of the dummy variables for the estimated changes in the actual stock midpoints are zero.

5.5. Signed volume during price disagreement events

We expect to see some "arbitrage" trading during disagreement events. For the {P > IP}-type disagreements, we expect that arbitrageurs buy call options at their ask prices and perhaps sell puts at their bid prices. For the {IP > P}-type disagreements, we expect that arbitrageurs buy put options and sell call options. In addition to arbitrage trading, there could be trading by patient option traders who follow the markets and wait for an opportune time to buy options. For example, a trader who wants to buy a call could recognize that a {P > IP}-type disagreement event is an advantageous time to buy. In looking for such trading, we measure option trading volume using the delta-equivalent share position; i.e., we compute the delta of each traded option, weight the deltas by the trade sizes, and sum the quantity-weighed deltas to obtain an estimate of the delta-equivalent share position traded in the options market.

In a frictionless market, we would also expect to observe trades in the underlying stock. But markets are not frictionless, and the evidence above indicates that the stock leg of the "arbitrage" trade will break even or perhaps lose money before transaction costs and, thus, lose money after transactions costs.[11] Thus, traders could elect to trade only in the options and not the stock, and no prediction is made about stock trading during the disagreement events. Because the disagreements tend to close quickly, little need exists to buy or sell stock to hedge any options trades, suggesting that there might be no unusual stock trading volume during the disagreement events.

For these reasons we focus on option trading volume during the disagreement events. We expect to see positive delta-equivalent volume during the {P > IP}-type disagreement events and negative delta-equivalent volume during the {IP > P}-type disagreement events.

[11] The estimates of the stock price movement indicate that the stock prices have a slight tendency to move so as to widen the disagreements, indicating the stock leg of the transaction tends to lose money.

Table 6
Average signed volume during the disagreement and control events.
Signed volume in the underlying stocks is based on the Lee and Ready (1991) algorithm, and a version of the quote rule is used to estimate the direction of options trades. The delta-equivalent volume is computed using the estimates of signed option volume and the option deltas computed from the Black-Scholes-Merton formula with implied volatilities from Option Metrics. "All options" include all option pairs for a given underlying stock, and "Disagreement pair" includes only volume in the option pair that triggered the disagreement event. The columns headed "Mean" and "Median" report the mean and median, respectively, of the 39 values for each of the stocks and exchange-traded funds (ETFs). The columns headed "4th", and "36th" report the 4th and 36th largest of the signed volumes for the 39 stocks and ETFs. The units are round lots of one hundred shares, so that for example "1" means one hundred shares. All results are reported for evaluation periods of both 30 and 60 seconds. The number of disagreements for each stock is reported in Table 1. Because each disagreement event is matched to three control events, the control samples are three times the sizes of the treatment samples.

	{P > IP}-type disagreements				{IP > P}-type disagreements			
Sample and evaluation period	Mean	Median	4th	36th	Mean	Median	4th	36th
All options								
30-second evaluation period								
Treatment	36.0	23.5	8.9	57.9	−32.8	−20.0	−53.9	−3.0
Control	8.4	4.1	−2.2	13.1	−11.5	−4.3	−20.8	−0.2
1 minute evaluation period								
Treatment	42.0	25.9	10.6	74.3	−42.7	−24.1	−70.9	−4.1
Control	10.4	5.2	−4.1	21.0	−14.6	−5.0	−23.1	−0.3
Disagreement pair								
30-second evaluation period								
Treatment	17.1	12.4	4.1	33.5	−15.3	−11.1	−29.0	−3.1
Control	1.9	1.2	0.2	3.9	−2.7	−1.7	−8.5	−0.3
1 minute evaluation period								
Treatment	19.5	14.3	4.6	35.5	−17.7	−13.8	−32.8	−2.6
Control	3.0	2.2	0.1	5.6	−2.8	−1.9	−6.2	−0.4
Stock								
30-second evaluation period								
Treatment	6.3	15.3	−140.0	72.8	−12.9	7.1	−152.0	84.6
Control	12.8	9.8	−40.6	63.3	0.2	1.1	−57.0	33.1
1 minute evaluation period								
Treatment	18.1	11.5	−196.8	209.8	−23.9	10.6	−197.7	121.5
Control	27.9	11.0	−73.2	118.0	8.5	7.6	−71.1	68.5
Stock in 2-minute pre-event period								
Treatment	249.1	117.2	−27.6	601.1	−141.4	−44.8	−600.6	60.1
Control	196.2	109.5	−2.1	567.3	−143.7	−75.6	−434.7	20.0

Table 6 presents summary statistics for the signed volume in the stocks and the delta-equivalent signed volume in the options for the treatment and control samples for evaluation periods of 30 and 60 seconds. The statistics are constructed in the following way. First, the mean signed volume and delta-equivalent option volume during the events are estimated for each stock. The table then reports the equally weighted mean and median of the stock-by-stock means, along with the 4th and 36th largest of the stock-by-stock means. The table reports the statistics for the signed volume in the underlying stock, delta-equivalent signed volume in all options on the underlying stock, and delta-equivalent signed volume in the pair of options that triggered the disagreement. The results for the {P > IP}-type events are shown in the left half of the table, and those for the {IP > P}-type events appear on the right.

Table 7 is similar to Table 6, except that it presents total instead of signed volumes.

The option volume follows the predictions above. For the {P > IP}-type disagreements, during the first 30 seconds following the beginning of the disagreement, options traders on average buy the equivalent of 36 hundred shares, which is much greater than the delta-equivalent option volume in the control sample. The difference in the median, 4th, and 36th highest delta-equivalent option volumes are also large. The table also shows the volumes only in the option pair that triggered the disagreement, indicating that about half of the option volume occurs in the call-put pair that triggered the disagreement. Comparing the volumes during the 30- and 60-second evaluation periods, for the treatment sample most of the option volume occurs during the first 30 second following the beginning of the disagreement. Comparing the signed volumes in Table 6 with the total volumes in Table 7, signed volumes are a large fraction of total volume, especially for the disagreement pair.

The results for the delta-equivalent signed option volume during the {IP > P}-type disagreements shown in the right-hand side of the table are similar, except that the signed volumes are negative instead of positive because for these events the option-implied price exceeds the actual stock price. Also, because the delta-equivalent signed volumes are negative the magnitudes of the volumes in the 4th highest volume stock exceed those in the 36th highest volume stock. Again, signed volumes are large fractions of total volumes, especially for the disagreement pair.

Turning to the signed volume in the stocks, for the {P > IP}-type disagreements, the mean (of the stock-by-stock means) signed volume is slightly lower in the treatment sample than in the control sample, but the median in

Table 7
Average total volume during the disagreement and control events.

The options volumes are delta-equivalent volumes computed using option deltas option deltas computed from the Black-Scholes-Merton formula with implied volatilities from Option Metrics. "All options" include all option pairs for a given underlying stock, and "Disagreement pair" includes only volume in the option pair that triggered the disagreement event. The columns headed "Mean" and "Median" report the mean and median, respectively, of the 39 values for each of the stocks and exchange-traded funds (ETFs). The columns headed "4th", and "36th" report the 4th and 36th largest of the signed volumes for the 39 stocks and ETFs. The units are round lots of one hundred shares, so that "1" means one hundred shares. All results are reported for evaluation periods of both 30 and 60 seconds. The number of disagreements for each stock is reported in Table 1. Because each primary sample disagreement event is matched to three control events, the control samples are three times the sizes of the treatment samples.

	{P > IP}-type disagreements				{IP > P}-type disagreements			
Sample and evaluation period	Mean	Median	4th	36th	Mean	Median	4th	36th
All options								
30-second evaluation period								
Treatment	104.4	61.0	30.5	194.9	105.2	64.5	28.5	162.6
Control	57.0	33.9	9.8	68.0	58.4	30.8	12.0	79.6
1 minute evaluation period								
Treatment	167.7	112.2	42.8	294.1	168.7	101.4	43.0	247.0
Control	105.8	59.9	20.4	128.4	106.0	55.8	22.0	149.8
Disagreement pair								
30-second evaluation period								
Treatment	29.8	21.0	9.8	61.6	30.3	20.7	9.1	63.1
Control	8.1	5.8	1.8	13.9	8.2	5.0	2.3	14.5
1 minute evaluation period								
Treatment	42.3	29.7	12.9	90.3	42.6	27.5	12.6	77.0
Control	14.9	10.1	3.5	26.0	14.4	9.8	3.8	23.9
Stock								
30-second evaluation period								
Treatment	653.7	373.7	161.3	1813.3	631.6	311.6	153.9	1687.5
Control	464.5	241.4	111.5	1276.1	460.2	220.3	116.6	1336.3
1 minute evaluation period								
Treatment	1188.6	659.5	289.9	3215.3	1137.1	593.5	293.6	2968.4
Control	870.6	480.6	207.1	2319.5	864.8	442.4	211.9	2479.5
Stock in 2-minute pre-event period								
Treatment	2039.1	1182.1	477.0	5101.9	1898.1	1077.4	477.0	4953.4
Control	1683.4	968.5	393.9	4600.2	1602.8	843.6	379.8	4345.3

the treatment sample is larger. The same is true for the {IP > P}-type events. In all cases, these mean and median signed volumes are small relative to the corresponding total volumes in Table 7. These results are not suggestive of significant arbitrage trading in the stocks during the disagreement events.

The bottom panels of Tables 6 and 7 show signed and total volume in the two-minute pre-event periods. Consistent with earlier results for pre-event returns, there is positive (negative) signed volume in the {P > IP} (IP > P)-type events, in both the treatment and control samples. This similarity between the treatment and control samples should be unsurprising, because the two-minute pre-event return was used in matching. Combined with the findings for the delta-equivalent option volume, these results are consistent with the hypothesis that disagreements are triggered by signed volume in the underlying stocks and then closed following delta-equivalent signed volume in the options.

6. Analysis of various subsamples

We next turn to the analysis of subsamples in which one might plausibly expect to find different results or in which one can be confident that the results are not due to differing latencies between the option and stock markets or other possible concerns about the quote data. We find similar results in these different subsamples, providing confidence that the results are a robust feature of the option market and are not driven by different latencies or other data issues.

6.1. *Subsamples in which one might expect to find different results*

The stock market initiates 88% of the disagreement events in the sense that, in 88% of the events, a stock market quote update is the proximate cause of the disagreement. It is reasonable to hypothesize that the 12% of events that are initiated by the options market might have different characteristics. Specifically, in these events it might be the case that the actual stock price moves toward the options market quotes, i.e., that price discovery occurs in the options market.

Column 2 of Table 8 shows the estimates of the coefficient on the disagreement dummy in regressions including the control variables explaining the implied bid, implied ask, and observed stock midpoint for the subset of disagreement events that are initiated by the options market. We obtain similar results in untabulated regressions that do not include the control variables. For comparison, the corresponding results for the full sample are shown in Column 1. In Panel A showing the results for the {P > IP}-type disagreements, the change in the implied ask is 4.5

Table 8
Conditional average quote changes in various subsamples.

Panels A and B show coefficient estimates for the {P > IP}- and {IP > P}-type disagreement events, respectively. In each panel, each column reports the coefficient estimates on the disagreement dummy and the associated *t*-statistics (in parentheses) for regressions explaining the option-implied bid, option-implied ask, and actual stock midpoint for the subsample identified in the column heading. In each subsample, all observations are pooled together and a single regression is estimated. Each column also reports the median disagreement duration, in seconds, and the number of treatment sample events. Because each disagreement event is matched with three control events, the total number of observations is four times the number of treatment events. The evaluation period is 1 minute, and the *t*-statistics are based on White heteroskedasticity-consistent standard errors.

Results for the full sample are provided in Column 1 for comparison. Column 2 headed "Option-initiated" contains the results for the subsample of events (and corresponding matched observations) triggered by changes in options quotes. Column 3 headed "Two pre-earnings days" is for the subsample of events that occur in the two days before an earnings announcement. Column 4 headed "Pre-event return > 0.3%" uses the subsample of events for which the absolute value of the 2-minute pre-event stock return exceeded 0.3%, and Column 5 headed "Year > 2004 " uses events from 2005 and 2006. Column 6 headed "Option volume > 80th percentile" is for the subsample of events that occurred on days in which option trading volume exceeded the 80th percentile of daily option trading volume for that underlying stock. Column 7 headed "Pre-event order imbalance" uses events in which the ratio of signed to total stock market volume in the two minutes preceding the event exceeds 0.5. Column 8 headed "Trade confirmed" is for the subsample of events with trade confirmed quotes, where trade-confirmed means that there were trades at both the call and put quotes used to compute the option-implied stock quote that causes the disagreement. Column 9 headed "Exchange confirmed" is for the subsample of events with exchange confirmed quotes, where exchange confirmed means that the call and put quotes used to compute the option-implied stock quote that causes the disagreement were quoted by at least two exchanges. Column 10 headed "Trade or exchange confirmed" uses the quotes that are either trade or exchange confirmed. Column 11 headed " > 10-second duration" includes only events in which the violating implied quote does not change in the first ten seconds after the disagreement is triggered.

Variable	Full sample (1)	Option-initiated (2)	Two pre-earnings days (3)	Pre-event return > 0.3% (4)	Year > 2004 (5)	Option volume > 80th percentile (6)	Pre-event order imbalance (7)	Trade confirmed (8)	Exchange confirmed (9)	Trade or exchange confirmed (10)	> Ten second duration (11)
Panel A: {P > IP}-type disagreements											
Implied bid	5.1	4.5	5.2	5.6	4.8	5.3	4.8	5.0	4.9	5.1	4.5
	(131.2)	(49.8)	(25.7)	(60.6)	(73.3)	(64.9)	(45.2)	(58.2)	(42.9)	(83.2)	(87.0)
Implied ask	5.9	6.8	6.5	6.8	5.9	6.5	5.7	4.9	4.9	5.5	4.1
	(134.8)	(60.5)	(27.7)	(67.6)	(78.4)	(69.7)	(44.7)	(48.7)	(38.8)	(79.9)	(66.0)
Actual stock midpoint	0.5	0.1	0.4	0.5	0.5	−0.1	0.6	0.8	0.2	0.6	0.3
	(14.5)	(−1.4)	(3.0)	(9.2)	(3.8)	(7.7)	(3.3)	(6.6)	(1.0)	(7.1)	(11.1)
Median duration	8.0	5.1	5.8	5.0	3.4	6.5	15.0	17.7	9.4	8.9	30.3
Number of treatment events	38,979	4,745	1,517	10,161	12,868	10,095	4,916	5,402	4,228	13,734	17,765
Panel B: {IP > P}-type disagreements											
Implied bid	−5.8	−6.5	−6.5	−6.4	−5.9	−6.0	−5.6	−4.9	−4.9	−5.4	−4.1
	(−140.4)	(−61.5)	(−24.8)	(−71.3)	(−84.4)	(−67.1)	(−39.3)	(−54.2)	(−40.8)	(−83.6)	(−70.4)
Implied ask	−5.0	−4.3	−5.5	−5.5	−4.8	−5.0	−4.8	−5.1	−5.0	−5.1	−4.5
	(−132.9)	(−42.7)	(−20.6)	(−64.7)	(−74.7)	(−57.6)	(−39.1)	(−62.2)	(−44.0)	(−82.3)	(−91.1)
Observed midpoint	−0.6	−0.4	−0.5	−0.6	−0.5	0.1	−0.5	−0.7	−0.2	−0.3	−0.4
	(−13.7)	(1.4)	(−1.7)	(−8.7)	(−2.9)	(−3.7)	(−3.4)	(−8.9)	(−3.7)	(−8.8)	(−13.2)
Median duration	7.9	5.1	6.1	5.1	3.2	6.4	16.0	16.8	9.4	9.1	30.6
Number of treatment events	42,045	4,714	1,487	11,344	13,832	10,706	3,807	6,154	4,714	15,141	19,045

cents, not much different from the change of 5.1 cents in the full sample. The change in the actual stock midpoint is −0.1 cents, which is close to zero and not significantly different from zero. In Panel B showing the results for the {IP > P}-type disagreements, the change in the implied bid is −6.5 cents, of greater magnitude than the −5.8 cent change for the full sample. The change in the actual stock midpoint is 0.1 cents, close to zero and not statistically different from zero. These results indicate that even when the disagreement is initiated by the options market, the options market moves in a direction that tends to close the disagreement, and there is only a very small change in the stock price. Results for signed volume reported in the Internet Appendix reveal the presence of signed option trading volume in the direction that tends to close the disagreements, though the magnitudes are not as large as in the full sample.

Earnings announcements are a time when important value-relevant information is often released. Perhaps the options market participates more in price discovery in the period immediately prior to earnings announcements?

Table 8, Column 3, shows the conditional average changes in the option-implied quotes and stock midpoint for the subset of disagreement events that occur in the two trading days prior to earnings announcements. For both the {P > IP}-type and {IP > P}-type disagreements, the changes in both the option-implied bid and ask are similar to but slightly larger than the corresponding changes in the full sample. The changes in the stock midpoint are also similar to those in the full sample. Recalling that in the full sample the stock price changes tend to increase rather than decrease the magnitude of the disagreement, this result for the stock price changes is evidence that the options markets do not participate in price discovery in the period immediately prior to earnings announcements. The point estimates for signed option volume for this subsample reported in the Internet Appendix are very similar to those for the full sample, though the *t*-statistics are smaller because of the smaller sample size.

We also consider the possibility that the markets' reactions to large stock returns differs from the reactions in the full sample, perhaps because large stock returns tend to be overreactions to information. If this is the case, one might expect to see stock price reversals in which stock prices move toward the unchanged or less changed option market quotes. This would be consistent with some price discovery occurring in the options markets.

Column 4 headed "Pre-event return $>0.3\%$" shows the conditional average changes in the implied bid, implied ask, and observed stock midpoint for the subset of disagreement events for which the return (for the $\{P>IP\}$-type events) or its negative (for the $\{IP>P\}$-type events) during the two minutes prior to the beginning of the disagreement event exceeded 0.3%. The results here are very similar to those in Column 3 for the two days prior to earnings announcements. For both disagreement types, the changes in both the option-implied quotes and stock price midpoint are similar to but larger than the corresponding changes in the full sample. The estimates for signed volume for this subsample are very similar to those for the full sample (see the Internet Appendix, available from https://www2.bc.edu/dmitriy-muravyev/PriceDiscoveryEquityOptions).

Options market liquidity increased during the sample period, suggesting that perhaps more price discovery might have occurred in the options market during the second half of the sample period. Column 5 of Table 8 addresses this hypothesis by presenting results for the subset of disagreement events that occurred during and after 2004. The average quote changes for this subsample are very close to the averages for the full sample shown in Column 1, providing no evidence of any change over time in the amount of price discovery occurring in the options market.

Using a VECM, Holowczak, Simaan, and Wu (2007) present evidence indicating that options quotes are more informative during periods in which there are either large numbers of options trades or significant signed order flow in the options market. Column 6 of Table 8 explores whether our methodology also finds a greater informational role of options quotes during periods of high option volume by looking at the subsample of price disagreement events that occurred on days in which option trading volume exceeded the 80th percentile of daily option trading volume for that underlying stock. Greater movement of the actual stock price quotes toward the option-implied quotes would indicate a greater informational role for the options quotes. Focusing on the rows in Panels A and B labeled "Stock" and comparing the results in Column 6 with the results for the full sample in Column 1, in both the $\{P>IP\}$ and $\{IP>P\}$-type disagreement events the point estimates in Column 6 are similar to those in Column 1 and continue to show a slight tendency for the stock price to move away from, rather than toward, the option quotes. Turning to the results for the options in the rows labeled "Implied bid" and "Implied ask," there is either identical or greater movement of the option quotes toward the actual stock price quotes than is found in the full-sample results reported in Column 1. At the risk of overinterpreting the point estimates, these results are consistent with the hypothesis that high option trading volume makes option quotes more informative in the sense that they catch up with stock price quotes more quickly, but not that they lead stock price quotes.

Harris (2003, p. 373) indicates that some arbitrage opportunities arise when liquidity demand in one market causes prices in that market to deviate from their fundamental values. This suggests the possibility that options quotes might be relatively more informative immediately following order imbalances in the stock market.[12] Column 7 of Table 8 explores this possibility by examining the subset of events for which the ratio of the signed stock market volume (or its negative, for the $\{IP>P\}$-type events) to total stock volume in the two minutes preceding the event exceeds 0.5. The results in both Panels A and B show that for this subsample the changes in the option-implied bid, option-implied ask, and actual stock midpoint are very similar to those in the full sample.

We conclude that the finding that the options market does not participate in price discovery about the level of the stock price is a robust one that is found in different subsamples in which it plausibly might not have been found. Given this, we turn to an examination of potential data concerns.

6.2. Are the quotes good?

As is often the case in the analysis of high-frequency data, data issues are a potential concern. In particular, we calculate price disagreement based on the reported NBBO at each moment, but are the reported quotes valid? Can option market participants trade at the reported quotes? While backing away from quotes is not a concern with fully electronic exchanges, the frequency of disagreement events is relatively low. It is conceivable that some fraction of the measured disagreement events could be the result of data or data transmission errors and that such data errors could impact the results.

Before turning to examining the subsamples of likely valid quotes are which option market participants can trade, it is worth emphasizing that the striking evidence of delta-equivalent signed option volume during the disagreement events is inconsistent with the hypothesis

[12] Order imbalances in the stock market might be symptomatic of information trade in the stock market, not liquidity demand. But conditional on the alternative hypothesis that option market quotes are informative, these could be events in which the option market quotes contain relatively more information.

that our results are due to quotes that are not valid because of data errors or other reasons.

We use two approaches to identity quotes that are highly likely to be valid. First, an option best bid (ask) is "trade confirmed" if there have been trades at exactly this price while the bid (ask) remains the best bid (ask), but not earlier than 200 seconds before the disagreement event. A problem arises because to trade confirm an option-implied stock quote, one should confirm both the call and put quotes that are inputs to the call-put parity relation. As trading is not frequent in options, this quote confirmation reduces the sample size by a factor of about seven. However, we do not need four options quotes to confirm both the option-implied ask and the option-implied bid, as only one quote is of interest when prices disagree. For example, for the {P > IP}-type it suffices to trade confirm only the option-implied ask quote.

A second way to verify quote validity is to consider only quotes displayed by multiple exchanges. For example, if both the International Securities Exchange (ISE) and CBOE report a bid of 1.15, it is more credible than if only the CBOE displays this price. We call an option-implied stock price "exchange confirmed" if both the call and put quotes used in the put-call parity relation are reported by at least two exchanges. As should be expected, this requirement considerably reduces the number of disagreement events that are identified. Option market makers have a strong incentive not to quote call and put prices that are inconsistent with the underlying stock price. The frequency with which market makers at two or more option market makers quote option prices inconsistent with the stock prices is much less than the frequency with which at least one options market maker does so.

Trade and exchange confirmations can be combined to increase the sample size. For example, a call price can be trade confirmed while a put price is exchange confirmed. We use this approach to increase the size of the sample of confirmed quotes.

Columns 8, 9, and 10 in Table 8 show results for the subsamples in which both the call and put prices used in the put-call parity are trade confirmed, both the call and put prices are exchange confirmed, and for the combined case in which the call and put prices are either trade confirmed or exchange confirmed. In each case, we also use only correspondingly confirmed control observations. The results are similar to those obtained with the full sample. Estimates of the stock price changes are close to zero, and the point estimates indicate that the stock price tends to move to slightly widen the disagreement, continuing to indicate that the stock price does not participate in price discovery.

6.3. *Differing latencies across markets*

We also entertain the hypothesis that there can be a systematic difference in the latencies in different markets. [13] Specifically, could we find that the options market follows the stock market simply because latency is possibly much larger in the options market? Three considerations make this seem unlikely. First, we are not aware of any systematic or anecdotal evidence that latencies across modern electronic exchanges differ by a lot. Second, we have no reason to suspect that our data provider introduced additional latency or time-stamp errors. The data provider time-stamped the different quotes as they hit the vendor's server, using a single clock. Moreover, our data provider's business consists of vending real-time data to various high-frequency trading desks and, thus, has strong incentives to minimize both latency and errors. Third, the hypothesis that the disagreement events are spuriously created due to differing latencies across the options and stock markets implies that we should not observe significant signed volume pressure in the direction that would eliminate the disagreements. Thus, the signed volume results are a compelling piece of evidence that our results are not spurious.

Nonetheless, we address this potential data problem by considering the subsample of treatment events in which the option-implied quotes that triggered the disagreement do not change for at least 10 seconds after the event is triggered, and we examine the quote changes over a 1-minute evaluation period. Differing latencies between the options and stock markets cannot drive the results for this subsample, because ten seconds exceeds any plausible difference in latencies.

The results for this subsample are shown in Column 11 headed " > 10 second duration." Similar to the results in the full sample, the coefficient estimate for the dummy variable in the regression predicting stock price changes indicates a slight tendency for the stock price to move to widen the disagreement, again providing evidence that the options market does not participate in price discovery. Also, similar to previous results, the option quotes move toward the stock price. However, in this subsample the magnitudes of the coefficients are not as large as in the full sample or most of the other subsamples. This result that the average movement in the option quotes is not as large in this subsample is unsurprising, because the criteria used to select this subsample was that the option quotes do not change for at least 10 seconds. Thus, this subsample is likely to be overrepresented with events from market conditions in which quotes are less likely to change.

7. Results for the Hasbrouck information share measure of price discovery

The Hasbrouck (1995) information share is a widely used measure of the proportion of price discovery that happens in each of two or more related markets. As discussed in Section 2, a literature uses this measure to study the proportion of equity price discovery that takes place in the option markets.

[13] We do not have to be concerned about the possibility that different exchanges use clocks that might not always agree. The data vendor time-stamps the different quotes as they hit the vendor's server, using a single clock. Thus, the only issue is the possibility of different latencies at different exchanges.

In the context of studying option market price discovery, the information share is based on a VECM of changes in two price series; the actual stock price and an option-implied stock price computed from option prices. We compute the option-implied stock price using the put-call parity relation and bid-ask midpoints. The use of bid-ask midpoints is standard in the information share literature, and our use of the put-call parity relation is consistent with our approach to analyzing the primary sample above, though in the analysis of the primary sample we do not use midpoints but rather take account of the bid-ask spread. Letting $\Delta p_t \in \Re^2$ be a vector of the changes in the two prices, as in Hasbrouck (1995) and Chakravarty, Gulen, and Mayhew (2004) the VECM has a moving average representation

$$\Delta p_t = B_0 \varepsilon_t + B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \cdots, \tag{12}$$

where $\varepsilon_t \in \Re^2$ is a vector of zero-mean innovations with covariance matrix Ω, $B_i \in \Re^{2\times 2}$ is a matrix of moving average coefficients, and $B_0 = I$, the identity matrix.

Table 9

Estimates of the upper and lower bounds of the Hasbrouck (1995) information share of the option market.

The second through seventh columns report summary statistics of day-by-day estimates of the upper and lower bounds of the option market information share. For each stock, estimates of the upper and lower bounds are computed each day. The table reports the number of daily estimates for each of the stocks and the means, medians, and standard deviations of the daily estimates for each of the stocks. The two rightmost columns report estimates based on a pooled sample combining data from all days for each of the stocks. The last three rows of the table report the means, minimums, and maximums of the statistics for the 39 stocks.

Stock ticker	Number of daily estimates	Summary statistics of day-by-day estimates of the Hasbrouck information share						Estimates based on pooled sample combining data from all days	
		Upper bound			Lower bound			Estimate of upper bound	Estimate of lower bound
		Mean	Median	Standard deviation	Mean	Median	Standard deviation		
AIG	813	6.2%	3.2%	8.7%	4.7%	2.2%	7.9%	10.1%	6.8%
AMAT	707	5.1%	1.9%	9.8%	4.8%	1.9%	9.6%	0.8%	0.3%
AMGN	829	4.4%	1.9%	7.8%	3.2%	1.2%	6.8%	2.2%	0.3%
AMR	762	8.4%	3.5%	14.3%	7.6%	2.8%	14.0%	0.5%	0.3%
AMZN	830	4.1%	1.7%	7.3%	3.4%	1.3%	6.7%	1.2%	0.0%
AOL	88	9.8%	3.6%	18.0%	9.9%	3.0%	18.3%	0.7%	0.5%
BMY	441	7.5%	3.2%	12.4%	6.8%	2.7%	11.9%	0.7%	0.4%
BRCM	856	3.8%	1.6%	6.8%	3.1%	1.3%	6.4%	0.7%	0.2%
C	344	6.9%	3.0%	10.7%	5.8%	2.3%	9.9%	3.5%	1.8%
COF	864	5.6%	2.4%	8.8%	4.3%	1.8%	7.7%	1.5%	0.5%
CPN	178	18.7%	5.4%	27.2%	18.5%	5.5%	27.1%	1.7%	1.3%
CSCO	726	6.2%	2.3%	11.7%	5.8%	2.4%	11.3%	1.0%	0.4%
DELL	731	4.4%	1.7%	8.9%	4.2%	1.5%	8.9%	1.0%	0.2%
DIA	297	8.0%	3.8%	10.4%	6.1%	2.9%	8.8%	1.3%	0.8%
EBAY	849	4.3%	2.0%	7.7%	3.0%	1.0%	6.8%	2.4%	0.4%
EMC	645	10.9%	4.3%	17.6%	10.5%	3.8%	17.5%	2.4%	1.6%
F	350	14.1%	5.3%	20.6%	13.9%	5.3%	20.3%	2.5%	1.9%
GE	457	5.9%	2.9%	8.2%	5.3%	2.4%	7.7%	3.6%	2.8%
GM	443	6.8%	2.7%	10.3%	6.1%	2.4%	9.7%	2.5%	1.1%
HD	676	5.9%	3.0%	9.1%	4.9%	2.2%	8.6%	1.1%	0.5%
IBM	806	6.5%	3.8%	9.1%	4.5%	2.1%	8.0%	0.8%	0.4%
INTC	781	4.6%	2.0%	7.2%	4.2%	1.9%	6.8%	0.8%	0.3%
JPM	346	6.1%	3.1%	8.8%	5.2%	2.1%	8.1%	3.6%	2.3%
KLAC	856	3.8%	1.7%	6.1%	2.6%	1.1%	5.5%	1.0%	0.1%
MMM	541	5.3%	2.9%	6.5%	3.7%	1.8%	5.2%	2.1%	0.6%
MO	470	6.8%	3.0%	10.6%	5.5%	2.1%	9.4%	3.9%	2.2%
MSFT	688	6.5%	2.8%	11.4%	6.1%	2.4%	11.0%	1.1%	0.6%
MWD	360	5.3%	3.2%	6.5%	3.8%	1.9%	5.3%	3.8%	2.0%
NXTL	443	5.4%	2.2%	10.8%	5.2%	1.8%	10.4%	1.2%	0.6%
ORCL	586	8.3%	2.4%	14.7%	8.1%	2.6%	14.3%	1.1%	0.5%
PFE	616	7.5%	3.3%	12.5%	6.8%	2.7%	12.1%	3.6%	2.4%
QCOM	807	4.2%	1.8%	7.6%	3.0%	1.2%	6.6%	2.2%	0.1%
QLGC	839	5.2%	2.2%	9.4%	4.1%	1.5%	8.6%	1.1%	0.2%
QQQ	352	10.0%	6.8%	10.5%	8.3%	5.3%	9.4%	2.8%	2.1%
QQQQ	282	5.5%	2.7%	7.1%	3.6%	1.8%	5.1%	2.0%	0.3%
SBC	199	6.7%	3.1%	10.3%	6.2%	2.7%	10.0%	2.4%	1.5%
SMH	816	7.4%	4.4%	9.4%	6.1%	3.2%	8.7%	3.4%	2.7%
TYC	641	6.8%	2.6%	11.8%	6.2%	2.4%	11.3%	1.3%	0.5%
XLNX	844	4.0%	1.6%	8.3%	3.4%	1.3%	7.9%	1.0%	0.2%
XOM	373	3.9%	2.0%	5.6%	3.1%	1.4%	5.0%	1.7%	0.7%
Mean	588.3	6.7%	2.9%	10.5%	5.8%	2.3%	9.9%	2.1%	1.1%
Minimum	88	3.8%	1.6%	5.6%	2.6%	1.0%	5.0%	0.5%	0.0%
Maximum	864	18.7%	6.8%	27.2%	18.5%	5.5%	27.1%	10.1%	6.8%

Letting B denote the limit of the sum of the moving average coefficient matrices

$$B = \lim_{k \to \infty} \sum_{i=1}^{k} B_i, \tag{13}$$

the two rows of B are identical. Letting the vector b be either of the rows, the variance of the permanent component of the quotes is $\sigma^2 = b\Omega b'$. If Ω is diagonal, the information share of the jth market is defined as

$$\mathrm{IS}_j = \frac{b_j^2 \Omega_{jj}}{\sigma^2}. \tag{14}$$

If Ω is not diagonal, the information share is not uniquely defined. Upper and lower bounds on the information share can be computed by considering Cholesky factorizations of all possible rotations of the innovations.

In estimating the VECM, we use price quotes observed at 1-second frequency and two hundred lags.[14] We include $k = 5000$ moving average matrices in the sum in Eq. (13). Even with five thousand terms, the two rows of the matrix B in Eq. (13) fail to agree in about 1% of the day-by-day estimations we describe below.[15] In these cases we do not compute the option market information share.

We compute two different sets of estimates of the upper and lower bounds on the option market information shares of the 39 stocks and ETFs. First, we follow the literature and for each stock compute estimates of the upper and lower bounds for each day. One reason to compute day-by-day estimates of the information share upper and lower bonds is that the VECM underlying the information share measure might not be stable over time. The second column of Table 9 reports, for each stock, the number of daily estimates that were computed. The numbers of daily estimates computed are not all equal to the maximum number of 864 because some stocks dropped from the sample before the end of the sample period, for some stock-days there is no call-put pair that passes the filters (a)–(d) described in Section 4.1, and for some stock-days the estimates of the rows of the matrix B do not converge even when including 5000 terms in the sum in Eq. (13). For each of the 39 stocks, the third and fourth columns of Table 9 report the means and medians, respectively, of these day-by-day estimates of the upper bounds, and the sixth and seventh columns report the means and medians of the day-by-day estimates of the lower bounds. The fifth and eighth columns report the standard deviations of the daily estimates for each of the stocks. The last three rows of the table report the means, minimums, and maximums across the 39 stocks of the various stock-by-stock statistics.

[14] After 200 lags, the VECM coefficients for the option market become statistically insignificant. For the stock market they go to zero even more quickly.

[15] With $k < 500$ we consistently obtain different estimates of the two rows of B. Our experience suggests that option market information share estimates computed from rows of a matrix B that have failed to converge are upward-biased. When the true information share is zero, any error, e.g., the use of rows of B that have failed to converge, results in an upward-biased estimate.

Second, we also pool the data for each stock and compute a single estimate of the upper bound and a single estimate of the lower bound for each of the 39 stocks. To avoid close-to-open jumps in stock prices, each day's stock prices were normalized so that the opening price was equal to the previous closing price. These estimates of the upper and lower bounds based on the pooled data for each stock are shown in the two rightmost columns of Table 9. A disadvantage of these estimates computed from the pooled data for each stock is that they involve the assumption that the VECM is stable over time. But if the VECM is stable over time, the information share estimates based on the pooled data will be closer to their probability limits. Because the information share is by construction between zero and one, when the true option market information share is zero estimates of it are upward biased. If the actual option market information shares are zero or close to zero, these estimates might be less biased than the means or medians of the day-by-day estimates.

The left-hand part of the table showing the results for the day-by-day estimates reveals that for most stocks the means of the day-by-day estimates of both the upper and lower bounds of the option market information share are in single digits, i.e., less than 10%. The medians of the day-by-day estimates are below the means for every stock and usually considerably below the means, consistent with right-skewed distributions of the day-by-day information share estimates for most of the stocks. Examining the last three rows of the table, the mean, minimum, and maximum of the stock-by-stock medians of the information share upper bounds are 2.9%, 1.6%, and 6.8%, respectively. The mean, minimum, and maximum of the stock-by-stock means of the information share upper bounds are larger, with the mean of the stock-by-stock means being 6.7% and the maximum being 18.7%. The statistics summarizing the estimates of the lower bounds are smaller than the corresponding statistics for the upper bounds.

These generally single-digit estimates of the option market information share obtained from day-by-day estimation of the upper and lower bounds are smaller than those reported in most of the literature. The estimates of the information share upper and lower bounds obtained by pooling the data for each stock are even lower. The means across stocks of the upper and lower bounds are 2.1% and 1.1%, respectively, and the maximums are 10.1% and 6.8%. While it is not clear whether one should prefer the day-by-day or pooled estimates, the two sets of estimates present a consistent picture of single-digit option market information shares. This evidence of some limited option market price discovery based on a statistical metric, the Hasbrouck information share, is consistent with the evidence of no economically significant option market price discovery reported in the main results of this paper, as it is not unusual for researchers to find evidence of statistically significant but economically insignificant predictability in financial markets. Our finding of single-digit option market information share also helps reconcile our results with those in the earlier literature, e.g., Chakravarty, Gulen, and Mayhew (2004). Option market price discovery appears to have declined since the sample period used in that paper. While the option market information share has not declined to zero, it has declined to a level that seems consistent with no economically significant option market price discovery.

8. Conclusion

This paper focuses on events in which the stock and options markets disagree about the stock price in the sense that the option-implied stock price obtained from the put-call parity relation differs from the actual stock price. During these disagreement events, the movement of the stock price is virtually identical to the movement of the stock price in otherwise similar events during which there is no disagreement about the underlying stock prices. Our findings about the magnitudes of the quote changes during disagreement events provide information about the economic significance of the statistical predictability found in the literature using VECMs and, therefore, complements this literature. Our results provide compelling evidence that option price quotes do not contain any economically significant information that has not already been reflected in the stock market price quotes and, thus, that in this sense option price quotes do not participate in the price discovery process for the underlying stock price. In contrast, during the price disagreement events, the option-implied stock price quotes move toward the actual stock price quotes.

We also use a much larger sample based on a weaker definition of a price disagreement and obtain similar results in that much larger sample.

In 88% of the primary sample disagreement events, the disagreement is precipitated by a movement in the underlying stock price quotes. During more than one-half of the events, there is signed order flow in the options market in the direction that will tend to push the option-implied stock price quotes in the direction of the actual stock prices. In some of the disagreement events, the signed order flow in the options market is large, consistent with some option market participants aggressively trading at option price quotes that have been rendered stale by the stock price movement.

Our main results are also found in the subsample of events for which the disagreement was precipitated by a change in an option price quote, subsamples in which the option price quotes are confirmed by either trades at the quotes or the presence of an identical quote at another exchange, and a subsample in which the results are almost certain not to be driven by differences in latency across the two markets.

References

Amin, K., Lee, C., 1997. Option trading, price discovery, and earnings news dissemination. Contemporary Accounting Research 14, 153–192.

Ansi, A., Ben Ouda, O., 2009. How option markets affect price discovery on the spot markets: a survey of the empirical literature and synthesis. International Journal of Business and Management 4, 155–169.

Anthony, J.H., 1988. The interrelation of stock and options market trading-volume data. Journal of Finance 43, 949–964.

Asquith, P., Pathak, P., Ritter, J., 2005. Short interest, institutional ownership, and stock returns. Journal of Financial Economics 78, 243–276.

Battalio, R., Schultz, P., 2006. Options and the bubble. Journal of Finance 61, 2071–2102.

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.

Blau, B.M., Wade, C. Comparing the information in short sales and put options. Review of Quantitative Finance and Accounting, forthcoming. Available at SSRN 1348133.

Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491–528.

Boehmer, E., Huszar, Z., Jordan, B., 2010. The good news in short interest. Journal of Financial Economics 96, 80–97.

Cao, C., Chen, Z., Griffin, J.M., 2005. Informational content of option volume prior to takeovers. Journal of Business 78, 1073–1109.

Chakravarty, S., Gulen, H., Mayhew, S., 2004. Informed trading in stock and option markets. Journal of Finance 59, 1235–1257.

Chan, K., Chung, Y.P., Fong, W.M., 2002. The informational role of stock and option volume. Review of Financial Studies 15, 1049–1075.

Chan, K., Kot, H.W., Ni, S.X., 2010. All are smart: do short sellers convey more information than option investors? Unpublished working paper. Hong Kong University of Science and Technology, Hong Kong.

Chan, K., Chung, Y.P., Johnson, H., 1993. Why option prices lag stock prices: trading-based explanation. Journal of Finance 48, 1957–1967.

Cremers, M., Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. Journal of Financial and Quantitative Analysis 45, 335–367.

Dechow, M., Hutton, A.P., Meulbroek, L., Sloan, R.G., 2001. Short-sellers, fundamental analysis, and stock returns. Journal of Financial Economics 61, 77–106.

De Jong, F., Donders, M., 1998. Intraday lead–lag relationships between the futures, options and stock market. European Finance Review 1, 337–359.

Desai, H., Ramesh, K., Thiagarajan, S., Balachandran, B.V., 2002. An investigation of the informational role of short interest in the Nasdaq market. Journal of Finance 57, 2263–2287.

Diether, B., Lee, K.H., Werner, I., 2009. Short-sale strategies and return predictability. Review of Financial Studies 22, 575–607.

Diltz, J.D., Kim, S., 2005. The relationship between stock and option price changes. Financial Review 31, 499–519.

Dong, W., Sinha, R.N., 2012. Where do Informed Traders Trade? Trading Around News on Dow 30 Options. Working Paper. University of Illinois, Chicago, IL (unpublished).

Easley, D., O'Hara, M., Srinivas, P., 1998. Option volume and stock prices: evidence on where informed traders trade. Journal of Finance 53, 431–465.

Fama, E.F., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.

Finucane, T.J., 1991. Put-call parity and expected returns. Journal of Financial and Quantitative Analysis 26, 445–457.

Finucane, T.J., 1999. A new measure of the direction and timing of information flow between markets. Journal of Financial Markets 2, 135–151.

Harris, F.H., McInish, T.H., Wood, R.A., 2002b. Security price adjustment across exchanges: an investigation of common factor components for Dow stocks. Journal of Financial Markets 5, 277–308.

Harris, F.H., McInish, T.H., Wood, R.A., 2002a. Common factor components versus information shares: a reply. Journal of Financial Markets 5, 341–348.

Harris, L., 2003. Trading and Exchanges: Market Microstructure for Practitioners. Oxford University Press, New York.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. Journal of Finance 50, 1175–1199.

Holowczak, R., Simaan, Y., Wu, L., 2007. Price discover in the US stock and stock options markets: a portfolio approach. Review of Derivatives Research 9, 37–65.

Imbens, G., Wooldridge, J., 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47, 5–86.

Jensen, M., 1978. Some anomalous evidence regarding market efficiency. Journal of Financial Economics 6, 95–101.

Jones, M., Lamont, O., 2002. Short-sale constraints and stock returns. Journal of Financial Economics 66, 207–239.

Lamont, O., Thaler, R., 2003a. Anomalies: the law of one price in financial markets. Journal of Economic Perspectives 17, 191–202.
Lamont, O., Thaler, R., 2003b. Can the market add and subtract? Mispricing in tech stock carve-outs. Journal of Political Economy 111, 227–268.
Lee, C., Ready, M.J., 1991. Inferring trade direction from intraday data. Journal of Finance 46, 733–746.
Mahalanobis, P.C., 1936. On the generalized distance in statistics. Proceedings of the National Institute of Sciences of India 2, 49–55.
Manaster, S., Rendleman Jr., R.J., 1982. Option prices as predictors of equilibrium stock prices. Journal of Finance 37, 1043–1057.
Nagel, S., 2005. Short sales, institutional investors and the cross-section of stock returns. Journal of Financial Economics 78, 277–309.
O'Connor, M.L., 2005. The cross-sectional relationship between trading costs and lead/lag effects in stock and option markets. Financial Review 34, 95–117.
Ofek, E., Richardson, M., Whitelaw, R.F., 2004. Limited arbitrage and short sales restrictions: evidence from the options markets. Journal of Financial Economics 74, 305–342.
Pan, J., Poteshman, A.M., 2006. The information in option volume for future stock prices. Review of Financial Studies 19, 871–908.
Savickas, R., Wilson, A.J., 2003. On inferring the direction of option trades. Journal of Financial and Quantitative Analysis 38, 881–902.
Stephan, J.A., Whaley, R.E., 1990. Intraday price change and trading volume relations in the stock and stock option markets. Journal of Finance 45, 191–220.
Vijh, A.M., 1990. Liquidity of the CBOE equity options. Journal of Finance 45, 1157–1179.