Journal of Empirical Finance 15 (2008) 839–849

Contents lists available at ScienceDirect

Journal of Empirical Finance

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

A comparison of trading and non-trading mechanisms for price discovery

Michael J. Barclay [1], Terrence Hendershott *

Haas School of Business, University of California, Berkeley, United States

ARTICLE INFO

Available online 18 March 2008

JEL classification:
G12
G14
G19
D82

Keywords:
Market microstructure
Price discovery
Price efficiency
Trading mechanisms
Volatility

ABSTRACT

This paper compares trading and non-trading mechanisms for price discovery during the Nasdaq pre-open and examines whether prices discovered though non-trading mechanisms are less efficient or reveal less information than prices discovered through trading. As Nasdaq pre-open trading volume increased, the opening price became more efficient and price discovery shifted from the opening trade to the pre-open. Price discovery shifted from the trading day to the pre-open only for the highest-volume stocks. These results suggest that pre-open trading contributes to the efficiency of the opening price, but that a critical threshold of trading volume is required to increase the amount of information in the opening price.

1. Introduction

The formation of prices in financial markets is a fundamental issue in financial economics. The revelation of private information through trading has been studied extensively in the microstructure literature (Kyle, 1985; Glosten and Milgrom, 1985; and Easley and O'Hara, 1987, and others). There are times during the day, however, when security prices must be discovered without the benefit of trading. A number of papers, for example, examine how security markets establish the opening price of the day (Stoll and Whaley, 1990; Biais et al., 1999; Cao et al., 2000; Domowitz and Madhavan, 2000; Madhavan and Panchapagesan, 2000; Davies, 2003; Barclay and Hendershott, 2003). Establishing an efficient opening price is important to financial markets. The open is the highest-volume period of the day and the opening price is used for a variety of off-exchange purposes such as the settlement of stock-index futures. However, discovering the opening price is a nontrivial undertaking. Both public and private information accumulate during the 17.5 h that the major U.S. stock exchanges are closed and the market impact of this information must somehow be determined.

The Nasdaq pre-open may be unique in the sense that there has been a dramatic shift in the behavior of market participants prior to the official opening of trade on the exchange. Cao et al. (2000) examined the Nasdaq pre-opening in 1994 when there was virtually no trading prior to Nasdaq's 9:30 a.m. opening. Although market makers were under no obligation to post quotes during the pre-open and any posted quotes were not binding, they observed that market makers routinely posted quotes in the pre-open and reacted to the quote changes posted by others. Cao, Ghysels and Hatheway argue that because informed traders could exploit their private information at market makers' expense, market makers had incentives to communicate with each other in an attempt

* Corresponding author.
E-mail address: hender@haas.berkeley.edu (T. Hendershott).

[1] A world-class scholar and gifted teacher, Michael J. Barclay, formerly of the Simon School of Business, University of Rochester, died in a seaplane accident on August 16, 2007. We thank Tim McCormick for providing data and the editor, two anonymous referees, and Mark Seasholes for helpful comments. Hendershott gratefully acknowledges support from the National Science Foundation.

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2008.03.001

to ensure that the opening bid and ask quotes reflected all available information. In the absence of trade, the method of communication was the posting of non-binding locked or crossed quotes.

During the following several years, Nasdaq stocks experienced an explosive growth of trading on electronic communications networks (ECNs). ECNs are electronic trading systems based on open limit order books where participants place orders and trade anonymously and directly with one another. This feature of ECNs facilitated the growth of after-hours trading. Because ECN trades do not require an intermediary, they have not been confined to exchange trading hours. As long as the system is turned on, trades can occur at any time of day or night. By 1999, spurred largely by the growth of trading on ECNs, the 250 most active Nasdaq stocks averaged more than a million dollars of pre-open trading per stock per day.

The growth of Nasdaq's pre-open trading allows a direct comparison of trading and non-trading mechanisms for price discovery. With no ability to trade during the pre-open in the early 1990s, Nasdaq market makers voluntarily signaled their private information with non-binding quotes. The advent of a functioning anonymous market for pre-open trading provided market makers with the ability to position their inventories and profit from their information through their trades. Thus, we compare Nasdaq opening prices in the non-trading regime of the early 1990s with opening prices discovered through trading in the late 1990s to provide new insights about trading and non-trading mechanisms for price discovery.[2] We also establish that TAQ is a reliable source for after-hours trades and can be used in future studies of after-hours trading.

We test two primary hypotheses about trading and non-trading mechanisms for price discovery. First, we test the hypothesis that trading in the pre-open allows market makers to establish opening prices that are less noisy and consequently more efficient than opening prices established with non-trading mechanisms.[3] Second, we test the hypothesis that trading in the pre-open increased the amount of information in the opening price by shifting the revelation of information from the trading day to the pre-open. If informed traders who previously waited until after 9:30 a.m. to execute their trades are now trading in the pre-open, then the information revealed through their trades would be shifted from the trading day to the pre-open.

We find that as the Nasdaq pre-open trading volume increased during the 1990s, the noisiness of the opening price declined, suggesting that pre-open trading has improved the efficiency of the opening price. However, for our sample of the 250 highest-volume stocks price discovery shifted from the opening trade of the day to the pre-open, but price discovery did not shift from the trading day to the pre-open.

Considering the emphasis in the literature that has been placed on trading in the price discovery process, it is surprising that for most stocks the opening price reflects the same amount of information with or without trading. Apparently, non-trading mechanisms effectively incorporated the impact of public information and revealed any private information that was highly depreciable. For example, market makers would like the opening price to reflect the information in their order imbalances because their inventories will move in the direction opposite from these trades. Without pre-open trading, market makers voluntarily signaled this information to the market through their non-binding quotes. With the advent of pre-open trading, market makers had a mechanism to reveal this information more credibly, more efficiently, and more profitably. However, the amount of information revealed was the same.[4] This supports the finding of Barclay et al. (1990) that a large volume of trade is required before private information is incorporated in prices. Our results suggest that informed traders with a low intraday rate of information decay delay their trades until uninformed traders are more active and trading costs are lower.

Larger trading costs in the pre-open – three to four times larger (Barclay and Hendershott, 2004) – deter pre-open trading and prevent price discovery from shifting from the trading day to the pre-open. While this is true for our entire sample, trading activity in the pre-open differs significantly across stocks: pre-open trading volume for the lowest-volume quintile is less than $50 thousand per day versus nearly $1.5 million per day for the highest-volume quintile. When examining price discovery by dollar-volume quintile, we find that the pre-open trading activity for the 50 highest-volume stocks is sufficient to move price discovery from the trading day into the pre-open, but the less active stocks have yet to reach the trading volume threshold to move price discovery from the trading day to the pre-open.

The remainder of the paper is organized as follows. Section 2 provides a general description of after-hours trading and compares the data available in TAQ with a complete record of after-hours trades from January through June of 1999. Section 3 examines changes in the efficiency of the opening price and after-hours trading from 1993 through 1999. Section 4 investigates how the changes in trading after hours impact the price discovery after hours and during the trading day. Section 5 concludes the paper.

2. Data and descriptive statistics

Two data sets are used for our analysis. The first, containing all after-hours trades of Nasdaq-listed stocks from January through June 1999, was obtained directly from Nasdaq.[5] For each after-hours trade, we have the ticker symbol, report and execution date and time, share volume, and price. The second data set is the TAQ database compiled by the NYSE.[6] TAQ data are used to obtain trades executed during the trading day (9:30 a.m. to 4:00 p.m.) and after-hours trades from 1993 through 1999. We compare the

[2] Cao et al. (2000) and Barclay and Hendershott (2003, 2004) characterize the Nasdaq pre-open during the non-trading and trading regimes, respectively. By examining both regimes with the same methodology, this paper is able to characterize the effects of trading on the price discovery process.

[3] Inefficiency in opening prices discovered through non-trading mechanisms could result from manipulation by large informed traders (Medrano and Vives, 2001).

[4] These results are also related to the broader literature on the relation between information, trading, and stock-return volatility. Return variances are higher during trading than non-trading hours (Fama, 1965 and others). French and Roll (1986) and Barclay et al. (1990) show that the higher stock-return variances are not fully explained by more public information released during exchange trading hours, but instead are caused by private information revealed through trade.

[5] We would like to thank Tim McCormick at Nasdaq for providing these data and facilitating our understanding of them and of after-hours activity in general.

[6] We filter out large data errors in both data sets by eliminating trades with large price changes that are immediately reversed and excluding trades with non-standard delivery options.

Table 1
Descriptive statistics for after-hours trading by time period

Time period	Average number of trades/day	Average volume/day ($000)	Average trade size ($000)	Median trade size ($000)
Pre-open (7:30 a.m.–9:30 am.)	6,471	341,858	53	28
Post-close (4 p.m.–6 p.m.)	9,428	964,253	102	32
Overnight (6 p.m.–7:30 am)	297	65,615	221	86
After-hours totals	16,196	1,371,726	85	30
Trading day totals	1,134,274	36,685,484	32	13
After-hours percent of trading day	1.43	3.74	261.86	242.25

The average daily number of trades, dollar volume, and average trade size for 3 after-hours time periods, the after-hours total, and the trading-day total for all Nasdaq stocks from January to June 1999.

TAQ data with the complete after-hours data obtained from the Nasdaq for the period during which they overlap. We then use the TAQ data to explore the evolution of after-hours trading over the 7-year period from 1993 to 1999. A detailed discussion of after-hours trading and trade reporting is available in Barclay and Hendershott (2003).

Table 1 characterizes the volume of after-hours trading in Nasdaq stocks in 3 after-hours time periods: the pre-open (7:30 to 9:30 a.m.), the post-close (4:00 to 6:00 p.m.), and overnight (6:00 p.m. to 7:30 a.m.). Nasdaq stocks collectively average 16,000 after-hours trades per day, totaling $1.4 billion or 3.74% of the average trading-day volume. After-hours trading in Nasdaq stocks is concentrated immediately after the close and immediately prior to the open of the normal trading day. Trading overnight is largely limited to late-night batch trading systems, the largest of which is Instinet's midnight crossing system.[7] This period also includes some trades between 6:00 p.m. and 7:00 p. m. or between 6:30 a.m. and 7:30 a.m. on high-volume days. After-hours trading volume is skewed toward the highest-volume days.[8] Twenty-three percent of the after-hours volume occurs on the busiest 5 days for each stock (of the 124 days in our sample period). Only 13% of the trading-day volume occurs on the busiest 5 trading days for each stock.

The mean and median trade size after hours is 2 to 3 times larger than during the trading day. The increase in the average trade size after hours is directly related to the lack of retail orders during this period. The largest after-hours trades occur in the post-close and overnight periods when the majority of trade prices are close to the 4:00 p.m. closing price.

2.1. After-hours trades in TAQ

If after-hours trading is to be widely studied by researchers, it is important to document availability and accuracy of public data. In TAQ, after-hours trades for Nasdaq stocks are marked with a "T" designator. Nasdaq trades are captured in TAQ when the Nasdaq Trade Dissemination Service (NTDS) is turned on, roughly 8:00 a.m. to 5:15 p.m. for the first 6 months of 1999. Given that after-hours trading volume is low outside of these times (Table 1), much of the after-hours activity may be available in TAQ. We explore this in more detail below.

In the complete after-hours data, there are 1,959,158 trades that occurred between 7:30 a.m. and 9:30 a.m. or between 4:00 p.m. and 6:00 p.m. During the same 6-month period, there are 1,935,251 "T" trades in TAQ (98.8%). The missing trades consist primarily of trades that occurred before the NTDS is turned on (from 7:30 a.m. to 8:00 a.m.) or after the NTDS was turned off (from 5:15 p.m. to 6:00 p.m.). In addition, there are approximately 40,000 trades that occur overnight, including the late-night crossing sessions, which are not in TAQ.

Fig. 1 examines whether the TAQ data accurately reflects the timing and distribution of trading volume during the pre-open and post-close. Fig. 1 plots the average daily trading volume in TAQ for each minute during the pre-open and post-close periods against the Nasdaq volume data. For each trade, the NASD data includes an execution time, in addition to the report time, which reflects a more accurate time stamp than is available in TAQ.

Fig. 1 shows that when the NTDS is active, the volume reported on TAQ and the volume reported by Nasdaq are nearly identical. If the number of trades is graphed, rather than dollar volume, the two lines are indistinguishable. Starting at about 4:30 p.m., the TAQ and Nasdaq volume data diverge slightly, and there are volume spikes in the Nasdaq data at 5:30 p.m. and 6:00 p.m. from a small number of very large trades with manually entered execution times of 5:30:00 p.m. and 6:00:00 p.m. During the pre-open, the TAQ report times and the NASD execution times are nearly identical because over 90% of the pre-open trades occur on ECNs (Barclay and Hendershott, 2003) that automatically report trades immediately.

There are relatively few after-hours trades that are not reported in TAQ, and the majority of missing trades occur overnight when there is little or no price discovery (Barclay and Hendershott, 2003). Therefore, an accurate time series of trading after hours can be constructed from the TAQ data. There were several important rule changes that affected after-hours trade reporting in 1993 and 1994.[9] The TAQ data may contain a lower fraction of all after-hours trades during these years. Although we use the TAQ data from 1993 to 1999 in the remainder of the paper, our results are not sensitive to the inclusion of 1993 and 1994.

[7] See Hendershott and Mendelson (2000) for details on the operations of crossing networks.
[8] At Nasdaq's request, the overnight period's beginning and ending times were chosen so it is not composed entirely of Instinet's midnight cross.
[9] During our sample period, the Nasdaq trade and quote reporting systems adhered to the following schedules (see Barclay and Hendershott 2003 for additional details):

June 15, 1992 Trade reporting required within 90 s for all trades executed from 9:30 a.m. to 5:15 p.m.
Dec. 20, 1993 90-s trade reporting requirement extended to 9:00 a.m. to 5:15 p.m.
Dec 12, 1994 90-s trade reporting requirement and ACT, NQDS, and NTDS extended to 8:00 a.m. to 5:15 p.m.
Oct 25, 1999 ACT, NQDS, NTDS and SelectNet extended to 8:00 a.m. to 6:30 p.m.
Nov 25, 1999 90-s trade reporting requirement extended to 8:00 a.m. to 6:30 p.m.

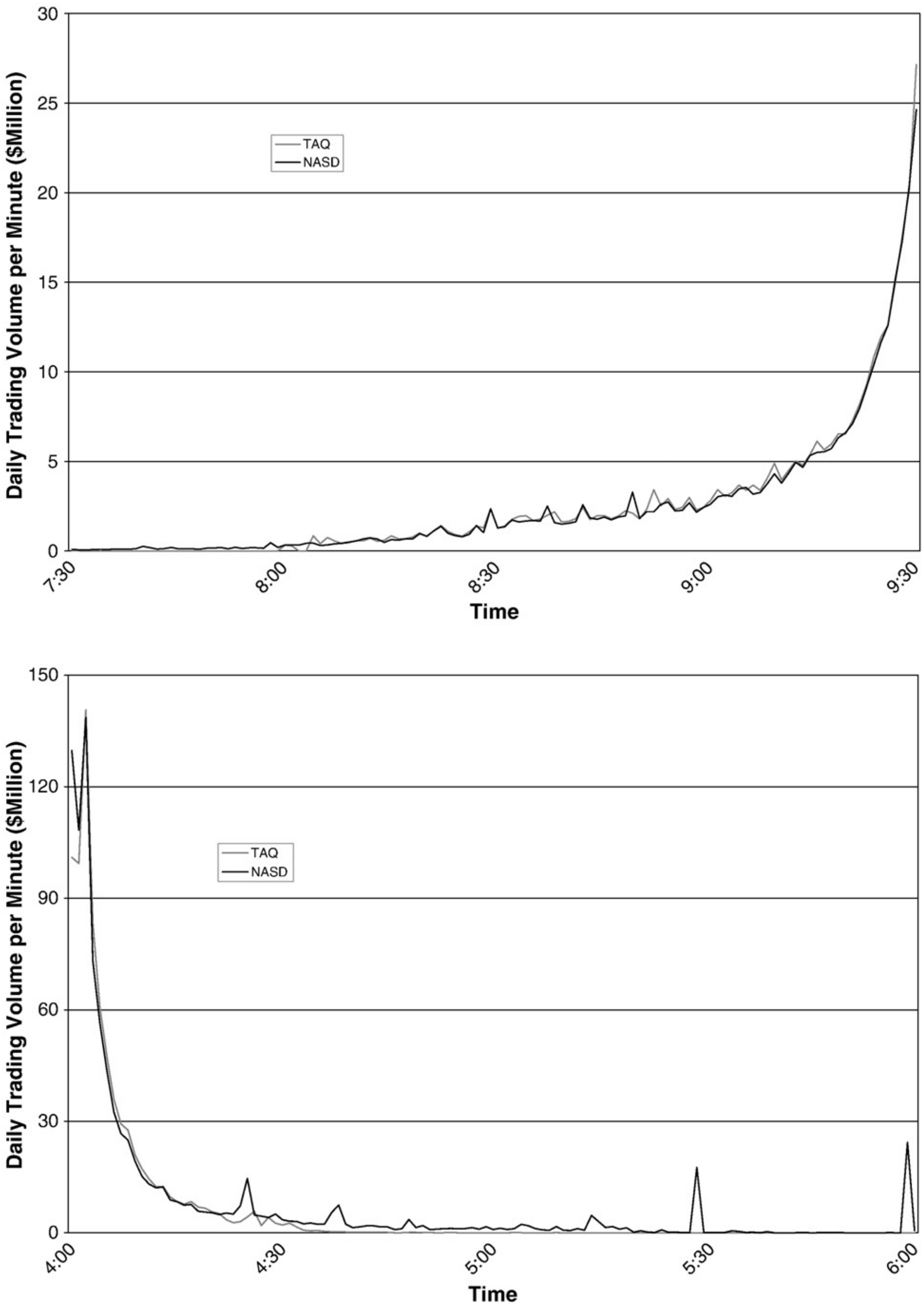

Fig. 1. Comparison of NASD and TAQ data for after-hours trades. Average daily trading volume by minute during the pre-open and post-close periods for the 250 highest-volume Nasdaq stocks from January to June 1999 using data from TAQ and the NASD.

Individual investors first entered the after-hours market in significant numbers in 2000. Because the after-hours participation of individual investors potentially confounds comparisons with earlier time periods, we end our analysis in 1999.

3. The efficiency of opening prices with and without trading in the Nasdaq pre-open

To examine the evolution of trading in the Nasdaq pre-open and its impact price efficiency and price discovery, we use TAQ to construct a sample of after-hours trades for the 250 highest-volume Nasdaq stocks (excluding ADRs) for each year from 1993

Table 2
After-hours trading for the 250 highest dollar-volume Nasdaq stocks by year

Year	Post-close			Pre-open			Trading day	
	Trading volume ($000)	Number of trades	Percent of days with trading	Trading volume ($000)	Number of trades	Percent of days with trading	Trading volume ($000)	Number of trades
1993	272	2.8	71.2	2	0.0	2.6	12,570	214.4
1994	357	3.5	76.7	35	0.4	15.7	13,686	255.2
1995	568	5.7	85.6	82	0.9	21.3	24,462	421.7
1996	686	7.0	89.4	202	2.0	39.7	31,687	577.8
1997	1,018	8.9	90.1	441	4.1	46.8	46,987	858.1
1998	1,311	11.3	91.9	611	7.1	64.4	61,618	1,327.1
1999	2,669	19.8	96.0	1,143	17.0	76.2	114,548	2,617.0

Average dollar volume and number of trades per stock per day, and percentage of days with at least one trade for the post-close, pre-open, and trading-day periods for each year for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ.

through 1999.[10] The 250 highest-dollar-volume Nasdaq stocks represent 75% of the after-hours trading volume and more than half of the after-hours trades for all Nasdaq stocks.

Table 2 provides details on the daily average trading activity for these stocks in the pre-open, post-close and during the trading day for each year from 1993 to 1999. During the trading day, the average trading volume increased by a factor of 10 over the 7 year period, from $12 million to $115 million per day, and the number of trades increased by a factor of 13, from 214 to 2617 trades per day. Pre-open trading volume was negligible in 1993 (approximately $2000 per stock per day) with pre-open trading on less than 3% of the trading days. By 1999, pre-open trading volume increased to $1.1 million per stock per day with pre-open trading on 76% of all trading days.

Fig. 2 plots the after-hours trading volume as a percentage of total volume by month. Except for the first 2 months of 1993, trading in the post-close increased at a rate similar to the trading day, remaining between 2 and 3.5% of total trading volume.[11] In contrast, trading in the pre-open was negligible in 1993 and has grown to approximately 1% of total trading volume in 1999. The increase in pre-open trading volume as a percentage of total volume is similar for both the high- and low-volume stocks in our sample.

3.1. The efficiency of the opening price

In this section, we examine whether trading in the pre-open has resulted in a more efficient opening price. Biais et al. (1999) propose a regression method, which they refer to as "unbiasedness regressions," to test the extent to which security prices reflect noise or information. We use these unbiasedness regressions to measure any changes during our sample period in the efficiency of the opening price.

For each stock each month, we regress the close-to-close return (ret_{cc}) on the return from the close to open (ret_{co}): $\text{ret}_{co}=\alpha+\beta\times\text{ret}_{co}+\varepsilon$. Although Biais, Hillion, and Spatt refer to these as "unbiasedness regressions," the slope coefficient (β) in these regressions has a natural interpretation as a signal-to-noise ratio. Consider the standard errors-in-variables problem for regression models (Maddala, 1988, p. 381).[12] If stock returns are serially uncorrelated and measured without error, then the slope coefficient in the unbiasedness regression would equal one. Suppose, however, that the "true" return process is unobserved, and that the observed return is equal to the true return plus noise. Noise in market prices can be related to microstructure effects (e.g., bid-ask spreads) or temporary pricing errors. In particular, suppose we observe $\text{ret}_{cc}=\text{RET}_{cc}+v$ and $\text{ret}_{co}=\text{RET}_{co}+u$, where RET_{cc} and RET_{co} are the "true" returns, and u and v have zero mean, and variances equal to σ_u^2 and σ_v^2, respectively. Then, regressing ret_{cc} on ret_{co} using ordinary least squares produces a slope coefficient, b, where

$$\text{plim}\ b=\beta\left(\frac{\sigma^2_{\text{RET}_{co}}}{\sigma^2_{\text{RET}_{co}}+\sigma^2_u}\right).$$

The term in parentheses can be viewed as the signal-to-noise ratio, where σ^2_{RETco} measures the information discovered from the close to the open, and σ_u^2 is the noise in the opening price. Although we cannot measure the signal and noise components separately with this technique, the extent to which b is less than one allows us to infer the signal-to-noise ratio.[13]

The unbiasedness regressions are estimated for each stock each month yielding a panel dataset consisting of 21,000 stock-month observations (250 firms spanning 84 months). Table 3 shows the average unbiasedness coefficient by year and trading volume quintile. The annual averages are calculated by first averaging the unbiasedness coefficient across stocks within a month and then averaging across months within a year. To account for the contemporaneous correlation across stocks induced by the common component in stock returns, we follow Fama-MacBeth (1973) and calculate the time-series standard error of the monthly mean.

[10] Sample firms are selected each year based on their total trading volume in that year. We require firms to trade in all 12 months of the year to enter our sample. Stocks ranked below 250 in total trading volume do not have enough pre-open trading activity for a meaningful analysis.
[11] It is possible that incomplete trade reporting during the first two months of 1993 caused the low reported post-close trading volume in that period.
[12] See Craig, Dravid, and Richardson (1995) and Barclay and Hendershott (2003) for related discussions.
[13] Actually, the measure is the ratio of signal to signal plus noise, but this terminology should not be confusing.

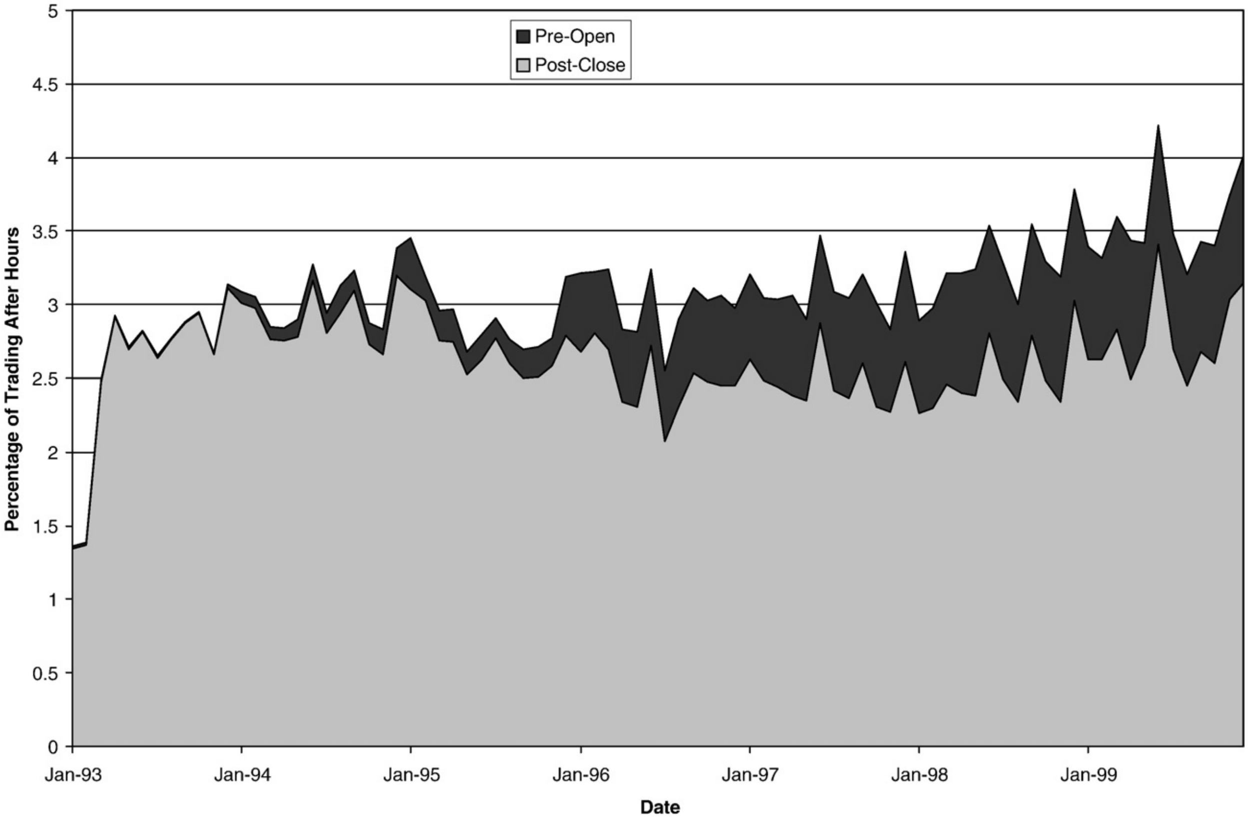

Fig. 2. After-hours trading volume as a percentage of total trading volume by month. The average daily dollar trading volume during the pre-open, post-close, and trading-day periods is calculated for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. The percentage of the total trading volume that occurs in the pre-open and post-close periods is provided in the graph.

Table 3 shows the efficiency of the opening price has increased over time, both for the entire sample of 250 stocks and for each trading volume quintile. Between 1993 and 1999, the overall opening price efficiency increases from 0.70 to 0.93. For the highest-volume quintile, opening price efficiency increases from 0.85 to 0.99 and for the lowest-volume quintile it increases from 0.63 to 0.85. The opening price efficiency is generally higher in the higher-volume quintiles.

Table 3
Opening price efficiency by year and trading volume quintile

	Volume quintile					
Year	Highest	2	3	4	Lowest	Overall
1993	0.851	0.682	0.705	0.637	0.628	0.701
	(0.040)	(0.038)	(0.031)	(0.030)	(0.032)	(0.045)
1994	0.870	0.702	0.712	0.646	0.656	0.717
	(0.040)	(0.025)	(0.026)	(0.036)	(0.017)	(0.037)
1995	0.900	0.850	0.768	0.751	0.680	0.790
	(0.079)	(0.068)	(0.049)	(0.039)	(0.045)	(0.060)
1996	0.919	0.822	0.833	0.779	0.755	0.822
	(0.063)	(0.053)	(0.038)	(0.041)	(0.026)	(0.047)
1997	0.952	0.893	0.821	0.767	0.762	0.839
	(0.097)	(0.061)	(0.054)	(0.052)	(0.035)	(0.065)
1998	0.876	0.901	0.881	0.813	0.792	0.853
	(0.064)	(0.055)	(0.052)	(0.051)	(0.046)	(0.053)
1999	0.986	0.974	0.936	0.924	0.849	0.934
	(0.066)	(0.078)	(0.048)	(0.040)	(0.036)	(0.059)

The close-to-close return is regressed on the return from close to open for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. Cross-sectional OLS regressions are run for each stock each month and the mean values of β from the regression $ret_{cc} = \alpha + \beta \times ret_{co} + \varepsilon$ is calculated across stocks each month. Days with zero price change are discarded. Standard errors are given in parentheses below and are based on the time series of monthly averages across stocks. Quintiles are constructed each year based on total dollar trading volume.

Table 4
Regressions of opening price efficiency on pre-open trading activity

Variable	(1)	(2)
Expected pre-open volume	22.078	
	(9.489)	
Unexpected pre-open volume	−12.923	
	(4.709)	
Expected pre-open volume*Q1		27.322
		(14.550)
Expected pre-open volume*Q2		28.948
		(13.939)
Expected pre-open volume*Q3		27.176
		(13.854)
Expected pre-open volume*Q4		18.193
		(11.972)
Expected pre-open volume*Q5		16.727
		(14.296)
Unexpected pre-open volume*Q1		−32.990
		(10.942)
Unexpected pre-open volume*Q2		−21.079
		(11.000)
Unexpected pre-open volume*Q3		−15.145
		(7.862)
Unexpected pre-open volume*Q4		−8.139
		(6.612)
Unexpected pre-open volume*Q5		2.368
		(5.314)
Log (mkt cap)	−0.098	−0.104
	(0.028)	(0.028)
Log ($ vol)	0.165	0.164
	(0.017)	(0.018)

The weighted price contribution for the close-to-close price change is divided into the open to close, close to last trade of the day, last trade of the previous day to 9:30 a.m., and 9:30 a.m. to the open for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. The opening price efficiency measure for each stock each month (β from the regression $\text{ret}_{cc}=\alpha+\beta\times\text{ret}_{co}+\varepsilon$) is regressed on the fraction of trading activity that occurs in the pre-open period (both expected and unexpected volume), the fraction of trading activity that occurs in the pre-open period interacted with dummy variables for the dollar volume quintiles (Q1 is the highest-volume quintile and Q5 is the lowest-volume quintile) and variables controlling for stock characteristics: log market capitalization (log(mkt cap)) and log daily dollar trading volume (log($ vol)). Firm fixed effects are included in all regressions, but not reported. The standard errors control for contemporaneous correlation and heteroskedasticity.

To examine the relation between pre-open trading and efficiency of the open more formally and to control for possible differences across stocks, we use the monthly stock estimates for the efficiency of the open in a panel regression. Firm-specific control variables – market capitalization (obtained from CRSP), stock-return volatility (standard deviation of daily close-to-close returns), and the log of average daily dollar trading volume – are calculated for each stock and each month.

Our primary focus is to determine how the increase in pre-open trading volume during our sample period has affected the efficiency of the opening price. First, the percentage of volume in the pre-open has increased during our sample period with a nearly linear trend. Second, pre-open volume is high for particular stocks on particular days when an information event occurs for that stock in the pre-open. For example, pre-open trading volume is high when firms announce important company news in the pre-open. To disentangle these two effects, we decompose the pre-open volume into its expected and unexpected components. Expected pre-open volume is defined as the fitted value from a regression of the percentage of volume in the pre-open on a linear time trend, the lagged value of pre-open volume and firm fixed effects. Unexpected pre-open volume is defined as the residual from this regression. This provides a good model of expected pre-open trading volume with an adjusted R^2 to about 81%.

To examine the relation between pre-open trading and the efficiency of the opening price, we regress the monthly unbiasedness coefficients on both expected and unexpected pre-open trading volume and firm-specific control variables, including firm fixed effects. The results are reported in Table 4. We use Rogers (clustered) standard errors throughout our analysis to control for contemporaneous correlation across stocks and heteroskedasticity (see Petersen, 2007) for a discussion of Rogers standard errors).

Consistent with the results in Table 3, regression (1) in Table 4 shows that the efficiency of the opening price increased as the expected trading volume increased from 1993 to 1999. To study whether the impact of pre-open trading volume on opening price efficiency is consistent across high- and low-volume stocks, we group our sample stocks into quintiles by total dollar trading volume and interact both expected and unexpected pre-open trading volume with dummy variables for each of the trading-volume quintiles. Quintile 1 contains the 50 highest-volume stocks and quintile 5 contains the 50 lowest-volume stocks in our sample. We then regress the unbiasedness coefficient on expected and unexpected pre-open trading volume interacted with the volume-quintile dummy variables and other firm-specific control variables. Regression (2) in Table 4 shows that the efficiency of the opening price increased as the expected trading volume increased for all five volume quintiles, although the effect is greatest in the higher-volume quintiles and not statistically significant in the lowest-volume quintiles.

Table 5
Weighted price contribution by time period and year

	Post-close	Pre-open	Open	Trading day
Year	Close-last trade	Last trade-9:30 a.m.	9:30 a.m.-open	Open–close
1993	0.043	0.007	0.118	0.831
	(0.003)	(0.001)	(0.005)	(0.005)
1994	0.046	0.033	0.086	0.834
	(0.002)	(0.002)	(0.004)	(0.005)
1995	0.034	0.044	0.062	0.860
	(0.003)	(0.002)	(0.005)	(0.006)
1996	0.030	0.060	0.051	0.858
	(0.002)	(0.003)	(0.005)	(0.006)
1997	0.018	0.074	0.045	0.864
	(0.002)	(0.005)	(0.006)	(0.009)
1998	0.013	0.094	0.044	0.849
	(0.002)	(0.009)	(0.004)	(0.011)
1999	0.011	0.118	0.018	0.853
	(0.001)	(0.007)	(0.002)	(0.008)

The weighted price contribution for the close-to-close price change is divided into the open to close, close to last trade of the day, last trade of the previous day to 9:30 a.m., and 9:30 a.m. to the open for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. The weighted price contribution is calculated for each stock for each month and then averaged across stocks and months. Standard errors are given in parentheses and are based on the time series of monthly cross-sectional averages.

Orders for Nasdaq securities that accumulate overnight are spread over many market makers, each of whom observe only a fragmented view of the order flow. Without trading, market makers employed a variety of non-trading mechanisms to signal this information to the market. The results in Table 4 indicate that these mechanisms produced an inefficient opening price. As pre-open trading volume increased, the efficiency of that price improves. The improvement over time in the efficiency of the opening price suggests that the pre-open trading has helped to generate a more efficient opening price and smooth the absorption of the overnight order imbalances.

While the trend of increasing pre-open volume has increased the efficiency of the opening price over time, an unexpected increase in pre-open trading volume reduces the efficiency of the opening price. As noted previously, high unexpected pre-open volume is likely to be associated with the arrival of information in the pre-open. The results in Table 4 indicate that the arrival of information and the associated increase in trading volume are associated with more noise in the opening price.

4. Price discovery with and without trading in the Nasdaq pre-open

The term "price discovery" is used to describe the process through which new information is incorporated in security prices. We extend our analysis of price discovery beyond the noisiness of price to examine the amount of new information that is reflected in pre-open and opening stock prices. We measure the information arrival in each time period by determining the fraction of the 24-h (close-to-close) stock return that is discovered in each period.

To quantify the amount of price discovery in each period, we use a measure called the weighted price contribution (WPC). The WPC measures the fraction of the close-to-close stock return that occurs in each period. We divide the 24-h day into four periods: the pre-open, the opening trade, the trading day and the post-close.

For each stock, and for each time period i, the WPC is defined as:

$$\mathrm{WPC}_i = \sum_{t=1}^{T}\left(\frac{|\mathrm{ret}_t|}{\sum_{t=1}^{T}|\mathrm{ret}_t|}\right)\times\left(\frac{\mathrm{ret}_{i,t}}{\mathrm{ret}_t}\right),$$

where $\mathrm{ret}_{i,t}$ is the log return during period i on day t and ret_t is the total close-to-close return on day t (from the close on day $t-1$ to the close on day t). The first term of the WPC is the weighting factor for each day. The second term is the relative contribution of the return for period i on day t to the total return on day t. The WPC normalizes the price discovery per period such that the WPCs sum to one.[14] As in the price efficiency analysis, we calculate the WPC for each stock each month to create a panel dataset that forms the basis for our analysis.

Table 5 provides the average WPC by time period (post-close, pre-open, open, and trading day) and by year. The annual averages are calculated by first averaging the WPC across stocks within a month and then averaging across months within a year. As with the price efficiency analysis we account for the contemporaneous correlation across stocks induced by the common component in stock returns by following Fama-MacBeth (1973) and calculate the time-series standard error of the monthly mean WPCs.

Table 5 shows that price discovery in the post-close is small throughout our sample period and decreases slightly from 4% in 1993 to about 1% in 1999. The consistently low and decreasing amount of price discovery in the post-close explains why most research on price discovery outside of the normal trading day has focused on the pre-open.

[14] Barclay and Warner (1993), Cao et al. (2000), and Huang (2002) also use WPC. The weights in the WPC reduce the heteroskedasticity in the observations and avoid the difficulties associated with zero price changes. Previous studies used price changes rather than returns. We use returns to make the results comparable across stocks and to facilitate the calculation of standard errors. The formulation is equivalent using price changes or returns. Log returns are used throughout the paper.

Table 6
Regressions of weighted price contribution on pre-open trading activity

	Pre-open		Open		Trading day	
	Last trade-9:30 a.m.		9:30 a.m.-open		Open–close	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Expected pre-open volume	18.231		−14.635		−2.003	
	(1.693)		(1.623)		(3.132)	
Unexpected pre-open volume	13.246		−0.466		−14.462	
	(0.845)		(0.612)		(0.945)	
Expected pre-open volume*Q1		23.233		−14.316		−15.160
		(2.482)		(2.059)		(2.786)
Expected pre-open volume*Q2		21.294		−16.463		−3.480
		(2.023)		(1.792)		(2.298)
Expected pre-open volume*Q3		15.908		−13.959		−0.425
		(1.842)		(1.705)		(2.458)
Expected pre-open volume*Q4		17.278		−13.758		−1.667
		(1.744)		(1.764)		(2.237)
Expected pre-open volume*Q5		15.213		−14.408		0.487
		(1.736)		(1.942)		(2.377)
Unexpected pre-open volume*Q1		22.351		−3.813		−24.129
		(1.944)		(1.391)		(2.289)
Unexpected pre-open volume*Q2		16.373		−0.710		−16.624
		(1.828)		(1.607)		(1.913)
Unexpected pre-open volume*Q3		12.532		−0.409		−13.117
		(1.268)		(1.392)		(1.307)
Unexpected pre-open volume*Q4		9.577		0.648		−10.715
		(1.412)		(1.181)		(1.424)
Unexpected pre-open volume*Q5		9.052		0.575		−10.878
		(1.432)		(1.230)		(1.674)
σ	0.917	0.929	0.168	0.167	−1.016	−1.034
	(0.152)	(0.154)	(0.076)	(0.076)	(0.161)	(0.162)
Log (mkt cap)	−0.021	−0.026	0.011	0.012	0.001	0.008
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
Log ($ vol)	0.030	0.029	−0.023	−0.023	0.011	0.013
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)

The weighted price contribution for the close-to-close price change is divided into the open to close, close to last trade of the day, last trade of the previous day to 9:30 a.m., and 9:30 a.m. to the open for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. For each time period weighted price contribution is calculated for each stock each month. The weighted price contribution is regressed on the fraction of trading activity that occurs in the pre-open period (both expected and unexpected volume), the fraction of trading activity that occurs in the pre-open period interacted with dummy variables for the dollar volume quintiles (Q1 is the highest-volume quintile and Q5 is the lowest-volume quintile) and variables controlling for stock characteristics: daily return standard deviation (σ), log market capitalization (log(mkt cap)), and log daily dollar trading volume (log($ vol)). Firm fixed effects are included in all regressions, but not reported. The standard errors control for contemporaneous correlation and heteroskedasticity.

The main shift that occurred over our sample period is a shift in price discovery from the opening trade to the pre-open. The sum of the WPCs for the open and pre-open periods remains between 11 and 14% throughout the sample period. However, price discovery in the opening trade falls from 11.8% in 1993 to 1.8% in 1999. Over the same period, price discovery in the pre-open rises from 1% to 11.8%.

The amount of price discovery during the trading day remained relatively constant over this 7-year period at about 85%. The small differences in open to close price discovery across years are not statistically significant.[15] It is somewhat puzzling that there was no shift in price discovery over our sample period from the trading day to the pre-open as pre-open trading volume increased (Fig. 2). It appears from these results that, although the non-trading mechanisms used to discover the opening price in the early 1990s produced inefficient opening prices (Tables 3 and 4), these mechanisms effectively determined the market effect of public information and discovered the highly depreciable private information. Market makers with private information about their accumulated net orders had incentives to signal this information to the market through their non-binding pre-open quotes. The lack of liquidity and high pre-open trading costs,[16] however, seem to have caused traders more durable private information to stay out of the pre-open and trade during the normal trading day. Thus, price discovery moves from the open to the pre-open as pre-open trading increases; however, the total amount of information reflected in the opening price remained unchanged.

To better examine the relation between pre-open trading and price discovery, we estimate panel regressions explaining the WPC in the pre-open, open, and trading-day periods. As in the price efficiency analysis, we decompose pre-open trading into its expected and unexpected components and include firm-specific control variables. The results of regressing the pre-open, open, and trading-day WPC on expected and unexpected pre-open trading volume and other firm-specific control variables are reported in Table 6.

[15] Barclay and Hendershott (2003) find that for 2000 the price discovery/WPC during the trading day is also 85% and the WPC at the open is essentially zero.
[16] Barclay and Hendershott (2004) find that in 2000 trading costs were 3 to 4 times higher in the pre-open than during the trading day, e.g., 9.1 cents during the trading day versus 38.7 cents in the pre-open. While quote data is unavailable to calculate the trading costs during our 1993 through 1999 sample period, the lower pre-open trading volume prior to 2000 suggests that the difference in trading costs between the pre-open and trading day are even greater during our sample period.

Regression (1) in Table 6 shows that increases in both expected and unexpected pre-open trading volumes are associated with significant increases in the pre-open WPC. Consistent with the results in Table 5, the positive and significant coefficient on expected pre-open trading volume indicates that as pre-open trading volume increased over the sample period, there was a significant increase in pre-open price discovery. Also, as expected, when pre-open trading volume was unexpectedly high, pre-open price discovery was also high.

Regressions (3) and (5) in Table 6 show expected pre-open trading volume having a negative and significant effect on the opening WPC, but not having a significant effect on the trading-day WPC. Also consistent with the results in Table 5, this indicates that as pre-open trading volume increased over the sample period, price discovery shifted from the opening price to the pre-open, but not from the trading day to the pre-open. Barclay et al. (1990) show that a significant amount of trading volume is required before private information is revealed through trading. Our results suggest that on average for our sample this volume threshold has not yet been reached in the pre-open. Although the increase in pre-open trading volume has sped up the effect of public information and highly depreciable private information (such as market makers' pre-open order imbalances) on stock prices, traders with more durable private information have not shifted their trades from the trading day to the pre-open. We would not expect this shift to occur until liquidity and trading costs in the pre-open are closer to those during the trading day.

In contrast to the expected pre-open trading volume, unexpected pre-open trading volume has a negative and significant effect on the trading-day WPC and an insignificant effect on the opening WPC. This is consistent with our earlier conjecture that pre-open trading volume is unexpectedly high on days with a pre-open information event. When information arrives in the pre-open that causes high pre-open trading volume, price changes in the pre-open will be larger than normal in relation to the trading day.

To determine whether the effect of pre-open trading volume is consistent across high- and low-volume stocks, we interact both expected and unexpected pre-open trading volume with dummy variables for each of the trading-volume quintiles. We then regress the pre-open, open, and trading-day WPC on expected and unexpected pre-open trading volume interacted with the volume-quintile dummy variables and other firm-specific control variables. The results are also reported in Table 6 in regressions (2), (4), and (6).

The regression results by volume quintile are generally consistent with those previously reported, although several differences emerge across volume quintiles. In regression (4) expected pre-open trading volume has a negative and significant effect on the opening WPC across all volume quintiles. Unexpected pre-open trading volume has a significant effect on the opening WPC in only the highest-volume quintile. In regression (6) unexpected pre-open trading volume has a negative and significant effect on the trading-day WPC for all quintiles.

In regression (2) increases in both expected and unexpected pre-open trading volume are associated with significant increases in the pre-open WPC for all volume quintiles. The size of the coefficients on expected pre-open trading volume declines from the highest- to the lowest-volume quintiles. This pattern likely reflects that fact that pre-open trading volume is measured as a percentage of total trading volume. The same increase in the percentage of volume in the pre-open represents a much larger number of trades and higher dollar volume for the high-volume stocks than for the low-volume stocks.

In regression (6) expected pre-open trading volume has an insignificant effect on the trading-day WPC for all but one of the volume quintiles. The lone exception to this pattern is the negative and significant coefficient for expected pre-open trading volume on the trading-day WPC for the highest-volume quintile. This suggests that the 50 highest-volume stocks have exceeded the pre-open volume threshold to move some price discovery from the trading day to the pre-open. To examine what this threshold may be for pre-open trading Table 7 provides per stock daily average dollar volume by year and dollar-volume quintile.

Panel A of Table 7 shows that while pre-open trading volume increases significantly in all quintiles, the volume increases between 1993 and 1999 are larger for the higher-volume quintiles. Over the entire sample the average daily trading volume for the highest-quintile stocks is $1.4 million. The second highest-quintile averages less than one eighth of that: $175 thousand. The fact that the coefficient on expected trading volume in regression (6) is significant for quintile 1 but insignificant for quintile 2 suggests that the pre-open trading volume threshold necessary to move some price discovery from the trading day to the pre-open lies somewhere between these values.

Table 7
After-hours trading by year and trading volume quintile

Year	Volume quintile					
	Highest	2	3	4	Lowest	Overall
1993	7	2	1	0	0	2
1994	145	15	6	7	3	35
1995	314	47	22	15	11	82
1996	703	141	69	58	39	202
1997	1,780	177	112	76	63	442
1998	2,418	273	157	119	87	611
1999	4,575	572	257	189	127	1,144
Overall	1,420	175	89	66	47	

Average dollar volume (in thousands) per stock per day for each year for the 250 highest-volume Nasdaq stocks from January 1993 to December 1999 using data from TAQ. Quintiles are constructed each year based on total dollar trading volume.

5. Summary

Establishing an efficient opening price is important to financial markets. However, determining the market impact of information arriving over the relatively long period the market is closed is a nontrivial undertaking. Different exchanges have developed different mechanisms to discover the opening price. The feature common to most of these mechanisms, however, is that the opening price is determined without the benefit of prior trading.

The Nasdaq pre-open may be unique in the sense that although there have been relatively few changes in the rules governing after-hours trades and quotes, there have been shifts in the behavior of market participants. In 1993, there was virtually no trading prior to Nasdaq's 9:30 a.m. opening. By 1999, the 250 most active Nasdaq stocks averaged more than a million dollars of pre-open trading per stock per day. The growth of Nasdaq's pre-open trading allows a comparison of the relative efficiency of trading and non-trading mechanisms for price discovery.

As the Nasdaq pre-open trading volume increased during the 1990s the noisiness of the opening price has declined over time, suggesting that pre-open trading has improved the efficiency of the opening price. In addition, much of the price discovery shifted from the opening trade of the day to the pre-open. Perhaps surprisingly, however, for only the 50 highest-volume stocks did price discovery shift from the trading day to the pre-open. This suggests that substantial amounts of pre-open trading is necessary to increase the amount of information in the opening price beyond what is possible using the non-trading mechanisms relied upon in the early 1990s. We also find that TAQ is a reliable source for after-hours trades and can be used in future studies of after-hours trading.

While a more efficient opening price is a public good creating value for market participants, the increase in pre-open trading has raised questions about whether or not market makers are treating their customers fairly and living up to their fiduciary duties (Ciccotello and Hatheway, 2000; Ip, 2000a,b, and others). The results in this paper suggest that the information that market makers may be trading on prior to the open, e.g., order imbalances, was bring incorporated into the opening price in the early 1990s when there was no trading before the open. This implies that any adverse price movements faced now by customers prior to the open would likely occur regardless of pre-open trading activity. Therefore, the only possible losers from the increase in pre-open trading would be liquidity providers facing high adverse selection. Barclay and Hendershott (2004) show that high trading costs in the pre-open just compensate liquidity providers for this high adverse selection. Thus, it seems likely that the increase in efficiency of the price open due to pre-open trading, rather than price signaling, comes at little cost to any market participants.

References

Barclay, Michael, Hendershott, Terrence, 2003. Price discovery and trading after hours. Review of Financial Studies 16, 1041–1073.
Barclay, Michael, Hendershott, Terrence, 2004. Liquidity externalities and adverse selection: evidence from trading after hours. Journal of Finance 59, 681–710.
Barclay, Michael, Warner, Jerold, 1993. Stealth trading and volatility: which trades move prices? Journal of Financial Economics 34, 281–305.
Barclay, Michael, Litzenberger, Robert H., Warner, Jerold B., 1990. Private information, trading volume and stock return variances. Review of Financial Studies 3, 233–253.
Biais, Bruno, Hillion, Pierre, Spatt, Chester, 1999. Price discovery and learning during the preopening period in the Paris Bourse. Journal of Political Economy 107, 1218–1248.
Cao, Charles, Ghysels, Eric, Hatheway, Frank, 2000. Price discovery without trading: evidence from the Nasdaq pre-opening. Journal of Finance 55, 1339–1365.
Ciccotello, Conrad, Hatheway, Frank, 2000. Indicating ahead: best execution and the Nasdaq preopening. Journal of Financial Intermediation 8, 184–212.
Craig, Alastair, Dravid, Ajay, Richardson, Matthew, 1995. Market efficiency around the clock: some supporting evidence using foreign-based derivatives. Journal of Financial Economics 39, 161–180.
Davies, Ryan, 2003. The Toronto Stock Exchange preopening session. Journal of Financial Markets 6, 491–516.
Domowitz, Ian, Madhavan, Ananth, 2000. Open sesame: alternative open algorithms in securities markets. In: Schwartz, Robert (Ed.), Building a Better Stock market: The Call Market Alternative. Kluwer Academic Publishing.
Easley, David, O'Hara, Maureen, 1987. Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69–90.
Fama, Eugene, 1965. The behavior of stock-market prices. Journal of Business 38, 34–105.
Fama, Eugene, MacBeth, James, 1973. Risk, return, and equilibrium—empirical tests. Journal of Political Economy 83, 607–636.
French, Kenneth, Roll, Richard, 1986. Stock return variances: the arrival of information and the reaction of traders. Journal of Financial Economics 17, 5–26.
Glosten, Lawrence, Milgrom, Paul, 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.
Hendershott, Terrence, Mendelson, Haim, 2000. Crossing networks and dealer markets: competition and performance. Journal of Finance 55, 2071–2115.
Huang, Roger, 2002. The quality of ECN and Nasdaq market maker quotes. Journal of Finance 57, 1285–1319.
Ip, Greg, 2000a. Role as big Nasdaq market makers helps Knight/Trimark's portfolio. Wall Street Journal March 3.
Ip, Greg, 2000b. Nasdaq looks to protect investors from dealers' pre-opening trades. Wall Street Journal May 3.
Kyle, Albert, 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Maddala, G.S., 1988. Introduction to Econometrics. Macmillan, New York.
Madhavan, Ananth, Panchapagesan, Venkatesh, 2000. Price discovery in auction markets: a look inside the black box. Review of Financial Studies 13, 627–658.
Medrano, Luis, Vives, Xavier, 2001. Strategic behavior and price discovery. RAND Journal of Economics 32, 221–248.
Petersen, Mitchell, 2007. Estimating standard errors in finance panel data sets: comparing approaches. Working paper.
Stoll, Hans, Whaley, Robert, 1990. Stock market structure and volatility. Review of Financial Studies 3, 37–71.