Public Information and Price Discovery: Trade Time vs Clock Time*

Shrihari Santosh§

(JOB MARKET PAPER)

November 2013

Abstract

I estimate the impulse response function of stock prices to news shocks using quarterly earnings surprises as an instrument for changes in firm value. I find that price discovery occurs through trading, independent of the passage of time. Prices converge to fundamental value faster when markets are more liquid. Nearly all cross-sectional variation in the speed of convergence measured in clock time disappears when measured in trade time. Price discovery requires minimal trading volume when liquidity is low, consistent with uninformed agents learning from order flow. I show that the path of asset prices in response to news is better explained by models of strategic informed trading than by gradual information diffusion.

*I thank my advisors Toby Moskowitz (Chair), John Cochrane, Bryan Kelly, and Juhani Linnainmaa for their guidance and inspiration. I am grateful for comments and suggestions from Torben Andersen, Ram Chivukula, Doug Diamond, Valentin Haddad, John Heaton, Yuan Hou, Serhiy Kozak, Damian Kozbur, Erik Loualiche, Ľuboš Pástor, Amit Seru, Wei Xiong, and participants at the Booth School of Business Faculty Seminar and the Economic Dynamics Working Group. Any errors are my own.

§University of Chicago Booth School of Business, ssantosh@chicagobooth.edu

1 Introduction

Financial markets function as mechanisms for information exchange as well as for capital allocation and risk sharing[1]. It is not well understood exactly how news comes to be reflected in prices. I find that over short time horizons, price discovery occurs through the process of trading itself. Kyle (1985) presents a model in which a strategic (non-competitive) informed investor spreads his orders over time to reduce the price impact of his trades. His measured trading results in prices only slowly impounding information as uninformed traders learn from order flow. With high-frequency transaction data, I show the path of prices in response to news is better explained by strategic informed trading than by the gradual diffusion of information across a population of competitive traders.

Using trade by trade data and quarterly earnings surprises as an instrument for long-horizon returns, I estimate cumulative impulse response functions (CIRF) of stock prices to information shocks. My procedure is similar to rescaling short-horizon cumulative returns by the ultimate long-run price impact of news. This extension of the event study method measures the fraction of information which is reflected in prices over time, permitting an apples-to-apples comparison across firms. I examine how well the estimated price paths fit the predictions of the strategic trading (Kyle 1985) and the gradual information diffusion (Hong & Stein 1999) models of price discovery.

Gradual information diffusion is the idea that knowledge spreads slowly across the population of traders and as this occurs, prices reflect more information. Asset prices do not fully impound the beliefs of the early informed traders due to limits to arbitrage and because uninformed traders do not learn from prices. Since there is no feedback from prices to beliefs, trading plays no role in price discovery. In contrast, the strategic trading model features uninformed agents who do learn through trading. Slow price discovery occurs because the non-competitive informed investor endogenously limits his trading intensity to prevent full information revelation. My empirical findings suggest that informed agents do behave strategically and uninformed agents do learn from market activity. Trading is important for the spread of information. Markets are not merely barometers reflecting agents' beliefs.

In my sample (2006-2011) earnings announcements typically occur during extended-hours trading sessions (*not* from 9:30 to 16:00). The official 9:30 opening price after the an-

[1]See Green (1973) for a model of information aggregation in asset markets. The basic idea that markets communicate information is developed by Hayek (1945), who notes that "we must look at the price system ... as a mechanism for communicating information if we want to understand its real function".

nouncement reflects 71% of the value shock indicating that significant price discovery occurs overnight. Within three minutes of regular trading the cumulative impulse response increases to 78% and by 10:30 it reaches 86%. Regular trading hours volume is an order of magnitude larger and bid/ask spreads are an order of magnitude smaller than during extended-hours trading (Barclay & Hendershott 2004). Back & Pedersen (1998), Collin-Dufresne & Fos (2012*b*) extend the Kyle model and show that the informed agent trades more aggressively and price discovery occurs faster when markets are more liquid[2]. The rapid increase in CIRF after the open is consistent with informed agents waiting for the predictable increase in liquidity. The gradual diffusion theory would attribute this pattern to a sudden increase in the rate of learning immediately after the regular trading session begins but is silent on why investors process information much faster in the few minutes after the market opens than at other times.

I decompose price discovery on the subsequent four days and find that nearly all convergence occurs during normal trading hours (9:30 to 16:00). Of the 10% remaining after event day, 8% occurs during regular trading hours and 2% during extended trading. This decomposition provides further evidence that prices impound news through trading and that price discovery occurs faster when markets are more liquid. To explain this pattern using a gradual diffusion model, uninformed agents must learn by significantly faster during the 6.5 hours of regular trading than during the remaining 17.5 hours of each day.

Using precise timestamps of earnings press releases and after hours transaction data, I measure price discovery trade by trade. For firms of all types, the CIRF jumps to only 20% at the first transaction post-announcement. The CIRF increases smoothly and rapidly to 65% by the thirtieth transaction. Median time elapsed since announcement to the thirtieth trade is only four minutes. As predicted by Holden & Subrahmanyam (1992), Foster & Viswanathan (1996) the CIRF is concave. In their models, concavity results from imperfect competition among multiple strategic informed traders. The gradual diffusion theory would imply that in the first four minutes post-announcement, 45% of traders analyze the news. During this time they consistently ignore the information content in prices[3].

To directly test the strategic trading and gradual information diffusion theories, I sort announcements into two types, *Fast* and *Slow*, based on the time elapsed from announcement to the thirtieth transaction. CIRF in trade time is nearly identical for the two groups, though they differ greatly in time elapsed. Median time taken for *Fast* and *Slow* is 70 and 950 seconds, respectively. This finding is consistent with the "market microstructure invariance"

[2] Liquidity is defined as variance of noise trader volume.

[3] Appendix A shows that the diffusion hypothesis relies on the assumption that early informed agents only employ "buy and hold" strategies. This results in time-inconsistency since in equilibrium they do rebalance. Allowing dynamic strategies delinks the CIRF from the spread of information.

hypothesis of Kyle & Obizhaeva (2011). They postulate that cross-sectional differences in market processes disappear "when trading...is scaled in units of business time instead of calendar time." I show that in high-frequency, "business time" is well approximated by the number of trades since announcement. CIRF measured in trades does not vary in the cross sections of share volume, dollar volume, and turnover.

After discussing related literature and data construction, Section 4 presents the empirical methodology and pooled estimation of impulse response functions. I use predictable variation in liquidity to test the hypothesis that price discovery occurs faster when markets are more liquid. In Section 5 I study cross-sectional variation in the speed of price discovery. I find support for the invariance hypothesis that apparent differences in the speed measured in clock time disappear when measured in trade time.

2 Related Literature

2.1 Price Discovery in Other Markets

There are a number of works documenting the process of price discovery in U.S. Treasury markets (cash and futures). Fleming & Remolona (1999) find that excess volatility and volume persist for ninety minutes after monthly CPI, PPI and employment releases. Beechey & Wright (2009) find that ten-year Treasury and TIPS yields fully respond to most macroeconomic news releases within ten minutes but that the information in FOMC announcements takes at least two hours to be fully impounded. The Treasury market is an order of magnitude more liquid than individual stocks, particular of small and mid-cap companies[4]. Thus, it is not surprising that public information is incorporated into Treasury prices more rapidly than into individual stock prices when measured in clock time. Green (2004) shows the asymmetric information component of bid/ask spreads on Treasury bonds increases after news releases which suggests "that market participants watch trading to help determine the effect of new ... information." Brandt & Kavajecz (2004) find that order flow imbalances result in permanent price changes even in the absence of macro news. Consistent with Kyle (1985), orders have larger price impact when liquidity is lower.

Andersen et al. (2007) find that stock, bond, and currency prices jump in response to news but they measure returns using five-minute returns and thus do not distinguish between

[4] Average daily notional volume of Treasury futures and options exceeds $200B on the CME alone. Effective spreads on the 10-year Treasury contract average around 2 b.p.

true jumps and continuous but rapid "drift." Consistent with uninformed agents learning through trading, Evans & Lyons (2002) find that daily movements in exchange rates and net order flow are strongly correlated (>0.7). Evans & Lyons (2008) estimate that roughly two-thirds of the impact of news on exchange rates occurs through order flow. They propose that "dealers observe macro news but have little idea how to interpret it, or how the rest of the market will interpret it. Instead, they wait to observe the trades induced and set their prices and expectations based on the interpretations embedded therein".

This paper also relates to the literature on disagreement, learning, return volatility and trading volume. Harris & Raviv (1993) and others show that differences of opinion (DO, or agreeing to disagree) generates significant trading volume around public information releases (consistent with data). Kim & Verrecchia (1994) propose a model of costly interpretation of public information. The model generates increased information asymmetry following a public announcement, leading to worsening liquidity and higher trading volume. In their model, the relatively less informed agents recognize the superior "quality" of the informed agents' analysis of public data. Banerjee (2011) derives predictions about the time-series characteristics of prices and disagreement and shows empirically that rational expectations (RE) learning by investors dominates the effect of agreeing to disagree. Carlin et al. (2012) find some empirical support for DO models but also find evidence that investors do "update their beliefs through a rational expectations mechanism when disagreement arises." The evidence in this paper suggests that RE and learning from trading contributes significantly to price discovery.

2.2 Rational Explanations of Predictability

One strand of the literature takes the position that post-event drifts reflect a priced risk which is not capture by popular models. Examples include liquidity risk (Sadka 2006), aggregate cash-flow risk (Ball et al. 2009), and disclosure risk (Shin 2006). Commonly cited evidence is the strong factor structure found in anomaly portfolios (Kim & Kim 2003, Chordia & Shivakumar 2006). I argue, however, that these particular risks are not likely to be large enough at high frequency to rationalize my empirical findings.

Shin (2006) develops a model where myopic managers "shade" the information in public disclosures to maximize the current stock price. This leads to increased short-run risk for positive surprises and decreased risk for negative surprises. With risk-averse investors this leads naturally to a risk-premium which looks like underreaction. In Shin's model, uncertainty is resolved through subsequent disclosures; hence, "drift" occurs at future disclosure dates (i.e. the subsequent earnings release). Over short horizons (minutes or days), manager myopia is probably not relevant.

Sadka (2006) argues that PEAD (post-earnings announcement drift) portfolio returns are correlated with shocks to systematic (market-wide) liquidity, thus earning a risk-premium. To rationalize my findings there must be "priced" high-frequency shocks to aggregate liquidity. Ball et al. (2009) claim that PEAD portfolios earn a risk-premium since returns are correlated with shocks to aggregate cash flow shocks. Again, to explain the predictable returns at the horizon I measure, returns immediately following an earnings release must be correlated with the aggregate earnings surprise of releases occurring in the next few minutes. Finally, some argue that observed price drifts are simply due to random chance (Fama 1998). The persistence of return predictability across time and markets suggests they are genuine phenomena and not statistical accidents.

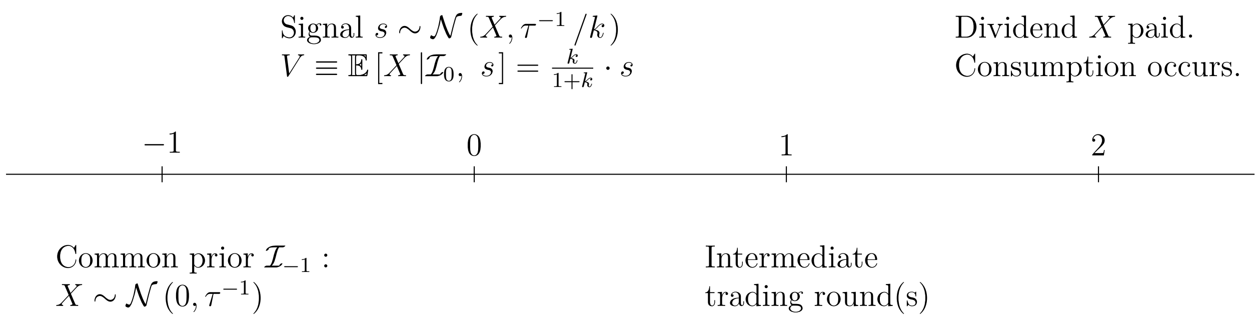
2.3 Behavioral Models of Price Discovery

Prior works have proposed various deviations from the rational, Bayesian benchmark to explain observed underreaction phenomena. The biases may be in preferences, such as prospect theory and disposition effect (Frazzini 2006), but more often are in the form of skewed learning. Examples are belief in wrong stochastic processes (Barberis et al. 1998), overconfidence in private information (Daniel et al. 1998, Kelley & Tetlock 2012), gradual information diffusion (Hong & Stein 1999), and inattention (Hirshleifer & Teoh 2003, DellaVigna & Pollet 2009). These theories depend on the assumption that "investors do not fully take into account the fact that they may be at an informational disadvantage, and hence do not draw the correct inferences from the trades of others" (Hong & Stein 2007); investors "agree to disagree."

The above models are intended to rationalize long-horizon (month to quarters) predictability. In this paper I argue for an alternative mechanism at short-horizons (minutes to days). I postulate market makers (liquidity providers) who do not condition their trading on all available information but who do learn from trading. A small number of traders correctly analyze public information and trade strategically to maximize profit. Why do the market makers ignore public signals? Heiner (1983) argues that uncertainty leads to rule-based decision making, which is the basis for predictable behavior. Faced with uncertainty regarding parameters, preferences, technology, etc, agents adopt relatively simple rules which seem to work well most of the time[5]. For market makers who regularly interact in markets with strategic, privately informed counterparties and who less frequently receive relevant public data, a rule derived from an adverse selection problem will be optimal most of the time.

[5] "...agents cannot decipher all of the complexity of the decision problems they face, which literally prevents them from selecting most preferred alternatives. Consequently, the flexibility of behavior is constrained to smaller behavioral reperotires that can be reliably administered."

Figure 1: Model Timeline



Only when there is unambiguous and freely available public information will the heuristic be inferior to full maximization[6]. Gabaix (2012) offers another formalization of imperfect maximization based on the principle of "sparsity", or using "few parameters that ... differ from the usual state of affairs." A main conclusion is that "when there is more uncertainty about the environment, the action is more conservative and closer to the default." For market makers, the default action is to set liquidity parameters based on the amount of adverse selection they face.

In Appendix A I solve competing models of price discovery in a common setting to generate predictions that depend only on the behavioral assumptions employed; Figure 1 shows the timeline. In each model I solve for the cumulative impulse response function

$$\theta_t \equiv \mathrm{CIRF}_t \quad \equiv \quad \mathbb{E}\left(\frac{P_t - P_{-1}}{V - P_{-1}} \middle| \mathcal{I}_{-1},\ s \right)$$

which maps directly to my empirical estimates. Below I summarize the intuition and main results.

Canonical inattention (or gradual diffusion) features a continuum of investors with exponential utility. By each time t a fraction μ_t of investors of investors have received the signal and update beliefs. The remaining $1 - \mu_t$ agents continue to rely on prior beliefs and do not learn from trading prices or quantities. Following Hong & Stein (2007), the μ_t informed agents form their asset demand using a static strategy where they intend on holding the asset until t_2 when the liquidating dividend is paid. With these assumptions, the equilibrium is given by $\theta_t \equiv \frac{(1+k)\mu_t}{1+k\mu_t}$. With a "weak signal" ($k \approx 0$), $\theta_t \approx \mu_t$; the cumulative impulse response at time t is equal to the fraction of agents who are informed by that date. This is the essence of the diffusion intuition. Taking the model to data, the estimated path of CIRF_t corresponds one-to-one with the diffusion of information.

[6] On the ex-date of a stock split, the first trade is on average unchanged from the last trade of the previous day after deflating by the split factor. This is an example of a large, salient, and obvious piece of information. It is no surprise that market makers fully react to such events.

I solve the same model but now allow informed agents to use dynamically consistent portfolio strategies. With this change, the tight link between θ_t and μ_t breaks down. The equilibrium is now $\theta_{-1} = 0$, $\theta_0 = \theta_1 = \frac{(1+k)\mu_1}{1+k\mu_1}V$, $\theta_2 = 1$. The key difference is that even if learning takes place from t_0 to t_1 ($\mu_1 > \mu_0$), prices do not change. Informed agents at t_0 anticipate that more investors will learn at t_1 and "frontrun" the soon to be informed. When informed use dynamic strategies, diffusion of information at time t does not generate "drift" at that date; the impulse response only increases at dates when informed agents face risk. Hence, large increases in CIRF_t must be accompanied by significant non-diversifiable risk.

Finally I solve a version of Kyle (1985) adapted to this setting. For analytical simplicity, I set $\mu_0 = \mu_1 = 0$ so that all competitive agents remain "uninformed". The model features a single risk-neutral informed agent who fully endogenizes the impact of his trading on equilibrium prices. In addition, at each date there is random inelastic "noise trading". Uninformed agents use a Kalman filter to extract the informed agent's beliefs from signed order flow. Because they are competitive, equilibrium satisfies $P_t = \mathbb{E}\left[\text{Firm Value}\,|\text{Prior beliefs \& trading history}\right]$; prices reflect the beliefs of competitive agents. Even without limits to arbitrage (risk aversion), imperfect competition leads the informed investor to optimally trade slowly to maximize his information rents, leading to slow price discovery. The model's solution is $\theta_0 \approx 1/3$, $\theta_1 \approx 2/3$. Convergenge to fundamental value occurs at t_1 even though no "diffusion" occurs at this date.

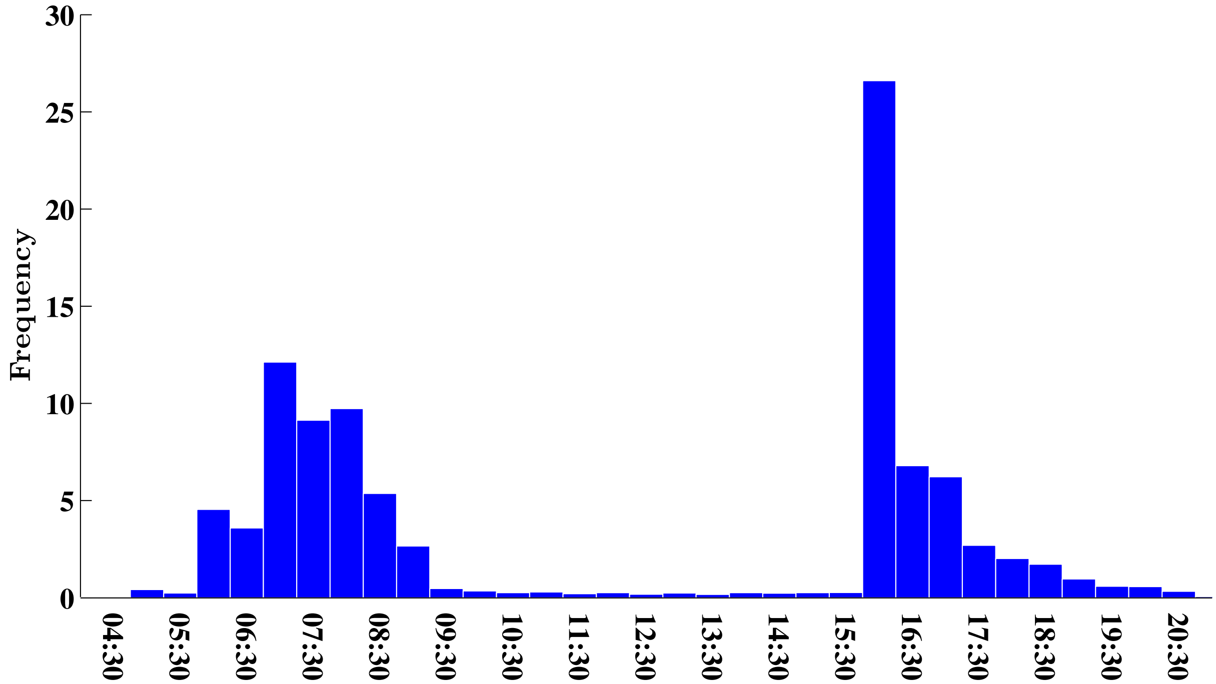
3 Data

3.1 Event Timing

Unlike previously studies, I rely on neither I/B/E/S nor Compustat for earnings announcement dates since these are known to contain date and time errors. From Wall Street Horizon (WSH), I get earnings press release dates and times, recorded to the nearest second[7]. WSH is a for-profit company exclusively devoted to accurate recording and prediction of earnings announcement, conference calls, and ex-dividend dates and times as well as EPS actual and estimated values. Figure 2 below shows the distribution of earnings announcement times (within day). The bimodal distribution has very small mass during trading hours. I exclude announcements that occur during the trading day since they may be substantially different along several unobserved dimensions.

[7] I manually verify 200 announcement times using Factiva and find 100% agreement between WSH data and the press release timestamp.

Figure 2: Histogram of Earnings Announcement Times

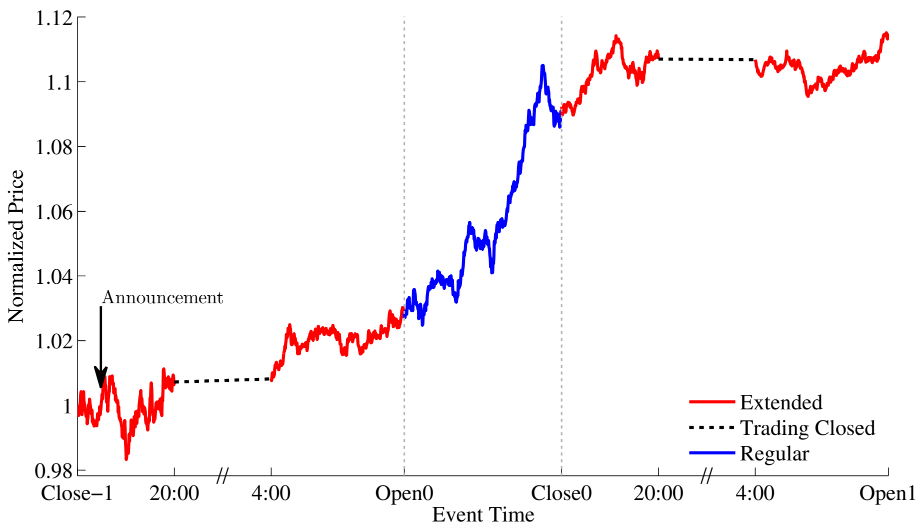


30 minute intervals. Time shown is (inclusive) left endpoint.

Figure 3 shows the event timeline for a hypothetical earnings announcement which occurs after market close (16:00) but before the next market open (9:30). Close$^{-}1$ represents the last closing price *before* the announcement. After-hours trading continues (potentially) until 20:00. Pre-market trading begin (potentially) at 4:00 the next day. Extended hours trading takes place on Nasdaq, NYSE Arca, and various other ECNs (electronic communication networks). Mechanically it is very similar to regular trading on the Nasdaq, where traders my both supply liquidity by posting limit orders or consume liquidity by executing immediately against an open order. A major regulatory difference is that the SEC's "Order Protection Rule" (Reg NMS Rule 611) does not apply during extended hours. Rule 611 integrates all exchanges during regular trading, essentially requiring that orders execute at the national best bid or best ask (best quote across exchanges). Pratically, extended hours trading is very different from the regular trading session. Extended hours bid/ask spreads are at least ten times larger and trading volume twenty times smaller relative to regular trading.

Open$^{+}0$ represents the first opening price *after* the announcement. NYSE and NASDAQ have similar auction mechanisms to set the official opening price as reported by CRSP. Close$^{+}0$ represents the closing price *after* one full regular trading session. Subsequent opening and closing prices are numbered sequentially. I use this convention for all announcements, whether they occur after market close (P.M.) or before market open (A.M.). The normalized price for firm i in quarter q at any horizon t, $P_t^{i,q}$, is the actual price, adjusted for splits and dividends, then scaled by the price as of $\text{Close}^{-}1$. This process normalizes the price as of $\text{Close}^{-}1$ to be 1 Let $B_t^{i,q}$ be the appropriate benchmark return for company i in quarter q. As commonly done, I define the abnormal return from event time to t as $CAR_t^{i,q} = P_t^{i,q}/B_t^{i,q} - 1$. In my baseline specification, I do not benchmark (set $B_t^{i,q} = 1$). This likely introduces no bias since I measure returns over short horizons during which risk premia are nearly zero and I include constants in regression specifications.

Figure 3: Event Timeline



I obtain opening and closing prices for individual stocks[8] from CRSP. I exclude returns based on bid/ask midpoint, keeping only actual prices from opening and closing auctions[9]. Intraday trade data is from NYSE TAQ, which includes all trades and quotes reported on the consolidated tape. For pre and after market trades, TAQ includes all transactions of at least 100 shares, but does not include trades on exchanges which do not participate in the National Market System. Barclay & Hendershott (2008) show that TAQ "is a reliable source for after-hours trades" though it technically does not contain the complete trading record[10].

Whereas nearly all equities (at least those which are covered by an institutional equity analyst) trade each day, the intraday trading record is more sparse. Instead of trade-by-trade analysis, I study intraday returns in constant clock time. From 9:30 to 16:00 I construct

[8]I include only firms listed on NYSE, Amex, NASDAQ, and NYSE Arca exchanges and securities with CRSP share codes 10 and 11 (Ordinary Common Shares with no further designation. This excludes ADRs, ETFs, REITs, and Closed-end mutual funds).

[9]Unreported results are similar using midpoints

[10]TAQ includes all trades which show up on the "Consolidated Tape" which covers all major U.S. exchanges and ECNs.

3-minute volume-weighted average prices (VWAP) for each firm using the trade data from TAQ [11]. Especially for smaller firms, this results in some missing "prices" throughout the trading day. Since I *do not* believe these data to be missing at random, I cannot simply ignore the missing observations. Instead of fully parametric maximum likelihood, I address the missing data problem using a version of random imputation [12]. To avoid extrapolation based on the imputation assumptions, I require non-missing opening and closing prices[13]. For analysis after hours, I measure price discovery in constant tick time, which does not suffer from issues of non-synchronous trading.

3.2 Earnings Surprise

My source of fundamental earnings data is Thomson-Reuters I/B/E/S from 2006 to 2011. For each quarterly earnings announcements of a U.S. firm, I construct $\widehat{EPS}$ as the mean of analyst forecasts in the inclusion window. I follow Thomson-Reuters and define the window as follows:

- For fiscal quarters 1, 2 and 3, the inclusion window starts 105 calendar days prior to the earnings announcement.
- For fiscal quarter 4, the inclusion window starts 120 days prior, since Q4 announcements typically occur more days after the end of the fiscal period. Since most estimates are revised shortly after an announcement, this extended window allows extra time for companies to report year-end results.

If an analyst issued multiple forecasts in the inclusion window, I use only the most recent one. I exclude announcements which occur more than 80 days after the end of the fiscal quarter (as recorded in I/B/E/S). I limit the sample to announcements that have valid stock return data in CRSP continuously for at least 20 trading days prior to the press release date. Furthermore, I exclude earnings announcements that are either less then 45 days or more than 150 days since the previous earnings announcement from the same company. This

[11] Following Huang & Stoll (1996), Bessembinder (1999), during regular market hours I include trades with *Sale Condition* of Blank, @, F, E, @F, and @E. This process excludes trades that are out of sequence, involve error corrections or non-standard settlement. For extended hours, I also include code T which indicates NASDAQ extended session trades. I further exclude transactions with zero price or volume. I exclude block transaction of >10,000 shares since they are likely to be executed in the "upstairs" market (crossing network), which may not be anonymous and is used by institutions for uninformed trading.

[12] To interpolate missing prices, I assume high-frequency returns are i.i.d normally distributed. Specifically, log prices follow a Brownian motion. Given two observations from a Brownian with known variance, the conditional normal distribution of intermediate values is given analytically (this is known as a Brownian bridge). Conditional means are derived from linear interpolation on the known values. Conditional variances follow a quadratic, with zeros at the two known points. Since I do not know the variance of the increments, I use estimate obtained from the cross-section of non-missing returns. The procedure draws random values from the distributions given by the analytic formulas.

[13] First and last 3-minute VWAPs

procedure excludes earnings events where the company did not report earnings the previous quarter or where the fiscal calendar changed, resulting in a significantly shortened quarter.

Following Kothari (2001), DellaVigna & Pollet (2009), I define the normalized earnings surprise as the difference between actual and forecast earnings per share (EPS), scaled by lagged stock price. Letting $EPS_{i,q}$ and $\widehat{EPS}_{i,q}$ represent the actual and forecast earnings per share for company i in quarter q and letting $P_{i,q}$ be the closing price per share of company $i's$ stock 5 trading days prior to the earnings announcement for quarter q, I define the normalized earnings surprise $S_{i,q}$ as:

$$S_{i,q} \quad = \quad \frac{EPS_{i,q} - \widehat{EPS}_{i,q}}{P_{i,q}}.$$

I exclude observations where $\|S_{i,q}\| > 0.01$ (which trims the data at the 5th and 95th percentiles). Finally, I exclude observation for which $P_{i,q} < 5$, eliminating penny stocks and potential data errors.

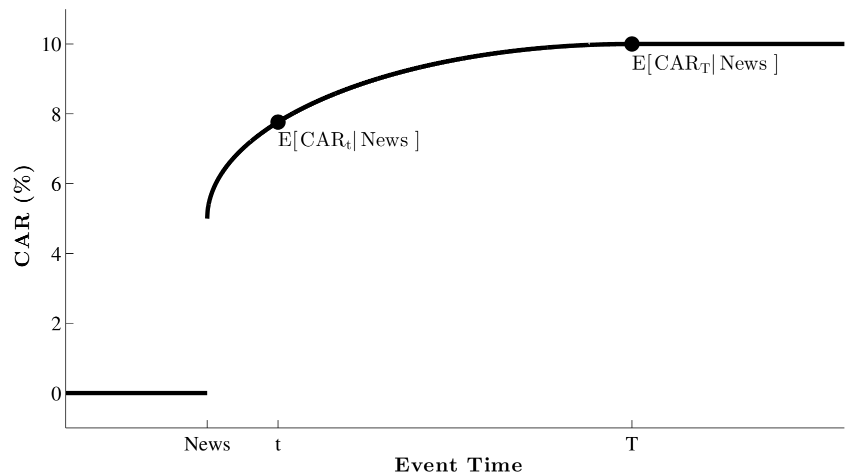
The earnings data from WSH includes company name and ticker symbol identifiers. Since CRSP includes both of these fields, I merge the datasets on ticker symbol and use the names to manually verify the integrity of each match. There are a non-trivial number of "duplicate" matches due to multiple securities trading under the same ticker, as reported tickers are usually truncated. When a firm has multiple share classes, the tickers are often of the form TICKER.CLASS. For example, Berkshire Hathaway has both BRK.A and BRK.B (the class B shares have a much lower price and no voting rights). Both WSH and CRSP truncate the class identifying portion of the ticker. When there are multiple matches, I keep the security with the lowest CUSIP number (7th and 8th digits). This identifies the primary share class for a firm. I merge the resulting dataset with I/B/E/S using PERMNO. I merge CRSP with TAQ using eight digit CUSIP.

After applying all filters I retain $\sim$ 35000 announcements from $\sim$ 3000 firms. Each quarter, the included firms represent approximately 70% of the total market value of CRSP universe.

4 Impulse Response

Event studies in finance typically ask two questions: (1) what is the average change in firm value given some event, $\mathbb{E}\left[\Delta \text{Value} \middle| \text{Event}\right]$, and (2) what is the path by which price reaches that value, $\mathbb{E}\left[\Delta \text{Price}_{\text{t}} \middle| \Delta \text{Value}\right]$. These manifest as an event-time graph of average CAR for the sample of firms which experienced the event as in Figure 4; the graph contains enough information to answer both questions. With a non-binary event, answering the questions requires specifying functional forms, or the common approach of comparing only the top and bottom deciles formed from a continuous measure.

Figure 4: Event Study



In this paper I estimate $\mathbb{E}\left[\Delta\text{Value}\middle|\,\text{Event}\right]$ flexibly but restrict $\mathbb{E}\left[\Delta\text{Price}_\text{t}\middle|\,\Delta\text{Value}\right]$ to be linear in ΔValue[14]. Assuming that prices embed news by time T, a natural measure of the speed of price discovery is $\theta_t \equiv {}^{\mathbb{E}[CAR_t]}/_{\mathbb{E}[CAR_T]}$ which measures the fraction of total price impact which has occurred by time t. Suppose θ_t does not depend on the size of the shock $\mathbb{E}\left[CAR_T\right]$. Then instead of restricting to binary news, θ_t can be estimated using all earnings surprises (or other news events). Furthermore, since θ_t is independent of the size of the shock $\mathbb{E}\left[CAR_T\right]$, it allows for an "apples-to-apples" cross-sectional comparison of speed. Finally, θ_t maps directly to the models in Appendix A which feature $CARA$ utility and normally distributed random variables.

I postulate the structural equations, omitting firm and quarter subscripts (i, q) from all variables unless otherwise specified:

$$\text{Normalize } P_{-1} \equiv P_{Close^-1} = 1 \tag{1}$$

$$\Delta V_t \equiv \mathbb{E}\left[P_t - P_{-1}\middle|\,\mathcal{I}_{-1},\ \text{News}\right] - \mathbb{E}\left[P_t - P_{-1}\middle|\,\mathcal{I}_{-1}\right]$$

$$= \mathbb{E}\left[P_t\middle|\,\mathcal{I}_{-1},\ \text{News}\right] - \mathbb{E}\left[P_t\middle|\,\mathcal{I}_{-1}\right]$$

$$V_t = \theta_t^{i,q} V$$

$$\theta_t^{i,q} = G_t\left(X_{i,q}\right) + \xi_{i,q,t}$$

$$\exists T < \infty,\ s.t.\ \forall i, q, t \geq T, \Delta V_t = \Delta V_T \equiv \Delta V$$

$$CAR_t \equiv P_t - P_{-1} = \theta_t \Delta V + \mathbb{E}\left[P_t - P_{-1}\middle|\,\mathcal{I}_{-1}\right] + e_t$$

$$\varepsilon_t \equiv \mathbb{E}\left[P_t - P_{-1}\middle|\,\mathcal{I}_{-1}\right] + e_T$$

[14]In Appendix D I provide evidence that linearity is a good approximation.

where $X_{i,q}$ is a set of observables and prices are deflated by the benchmark as specified in Section 3.1. All expectations are taken under the objective probability distribution using all public and private information. ΔV_t is the change in expected price at horizon t. Equation 1 specifies that news is incorporated into prices in finite time since returns are not predictable beyond horizon $T < \infty$. ΔV is the change in "long-horizon" prices but can be interpreted as a change in fundamental value. I also specify a linearity restriction: changes in expected prices at any horizon are linear in the long-run shock. This means that $\theta_t^{i,q}$ does not depend on the size or sign of ΔV. Throughout this paper $G_t(X_{i,q})$ takes various context dependent forms. By iterated expectations, $\mathrm{Cov}(\Delta \mathrm{V_t},\ \varepsilon_\mathrm{t}) = 0$.

(3) a function of the earnings surprise, $z(S_{i,q},\ X_{i,q})$, is a valid proxy for the change in firm value $\Delta V_{i,q}$; $\mathbb{E}[CAR_t|\,\Delta V,\ z(S_{i,q},\ X_{i,q})] = \mathbb{E}[CAR_t|\,\Delta V]$.

4.1 Pooled Estimation

In this section $G_t(X_{i,q})$ is simply a constant κ_t. I use all available observations to estimate a single θ_t for each horizon t. In addition to the structure specified in Equation 1, I further require:

$$z(S) \in \text{News}$$
$$\mathrm{Cov}(\Delta V, z(S)) \neq 0$$
$$\forall t,\ 0 = \mathrm{Cov}(\mathbb{E}[P_t - P_{-1}|\,\mathcal{I}_{-1}], z(S))$$

where S is the scaled earnings surprise and $z(S)$ is a function specified by the econometrician. The above conditions mean that returns are forecastable using $z(S)$ and that it is uncorrelated with prior expected drift, $\mathbb{E}[P_t - P_{-1}|\,\mathcal{I}_{-1}]$. Given the assumptions and iterated expectations, $\mathrm{Cov}(z(S), \varepsilon_t) = 0$. This implies that $z(S)$ is a valid proxy for ΔV, $\mathbb{E}[CAR_t|\,\Delta V,\ z(S)] = \mathbb{E}[CAR_t|\,\Delta V]$.

I specify $z(S)$ as a cubic polynomial with estimated coefficients, but my results are robust to using a linear specification, inclusion of higher order terms, or sorting S into bins and using a vector of indicator variables (see Appendix D). Later I explicitly model heterogeneity by allowing $G_t(X_{i,q})$ to vary with observable firm characteristics. I estimate θ_t at each horizon by two-stage least-squares:

$$\begin{aligned} \text{2nd Stage}: \mathbb{E}[CAR_t|\,S] &= a_t + \theta_t \cdot \widehat{\mathbb{E}}[CAR_T|\,S] \\ \text{1st Stage}: \mathbb{E}[CAR_T|\,S] &= \alpha + \gamma_1 S + \gamma_2 S^2 + \gamma_3 S^3 \end{aligned} \tag{2}$$

which recovers a conditional-variance-weighted average of the individual firm $\theta_t^{i,q}$[15]. Using

[15]Equation 1 explicitly acknowledges unobserved heterogeneity ($\xi_{i,q,t}$), meaning that two-stage least

an instrument is necessary since the second stage alone suffers from classic measurement error as well as omitted variables bias. In the first stage, I use the first three Hermite polynomials in S, adapted to the sample mean and variance. For any given degree, using Hermite or standard polynomial terms generates identical second stage estimates. Using Hermite polynomials, however, assists in comparing first-stage using different degrees. This is because when adding terms, coefficients and standard errors for lower order polynomials will be unchanged when using an orthogonal basis (like Hermite polynomials) but not when using standard polynomials due to multicollinearity.

4.1.1 Choosing T

To implement the procedure specified by Equation 2 I must specify T. The choice is not innocuous since by assumption all relevant information from the event is impounded into prices by this time. Table 1 analyzes various choices of post-event horizon (first-stage dependent variable). Each row represents a different T, ranging from 1 to 30 days since the pre-event close. Each column represents a choice of t for the second-stage dependent variable CAR_t; reported coefficients are θ_t. Going down each column, the estimates decrease monotonically until the sixth row ($Close^{+}5$) at which point they stabilize. This means there is no "drift" after six trading days post-event. The t-statistics (in parentheses) also decrease down each column since longer horizon returns naturally have higher variance resulting in less precise estimates. Lack of drift after one week contrasts with prior studies which find significant predictable returns for months after earnings announcements. The difference could come from sample selection since I estimate using more liquid stocks. It may also arise from data timing; I estimate using earnings from 2006-2011 whereas earlier studies typically use data from 1990-2005. I apply the methodology in DellaVigna & Pollet (2009) using data from 2006-2011 and find drift for at least one month. After trimming the data at the 5th and 9th percentiles of the surprise distribution I find predictability for only one week, consistent with Table 1. Kim & Kim (2003) find that drift after two days is absorbed by a PEAD factor. I conclude that long-horizon drift is a phenomenon isolated to very extreme surprises, which occur primarily for very small firms. I choose ($Close^{+}5$) as T since it appears to be the shortest horizon at which θ_t is no longer changing. Using any shorter horizon produces estimates which are biased upward. Using any longer horizon results in less precise estimates of θ_t.

squares recovers a conditional variance weighted average instead of a simple average slope (Angrist 1998). See Appendix C for properties of the estimator.

Table 1: Cumulative Impulse Responses

$$\begin{aligned} \text{2nd Stage}: \mathbb{E}\left[CAR_t \mid S\right] &= a_t + \theta_t \cdot \mathbb{E}\left[CAR_T \mid S\right] \\ \text{1st Stage}: \mathbb{E}\left[CAR_T \mid S\right] &= \alpha + \gamma_1 S + \gamma_2 S^2 + \gamma_3 S^3 \end{aligned}$$

Each row represents a choice of T and each column a different horizon t. Estimates are cumulative impulse response at horizon t. Standard errors are adjusted for heteroskedasticity and are clustered by firm and month of announcement. Absolute t statistics in parentheses.

$T \backslash t$	Open$^+$0	9:30	10:00	15:00
Close$^+$0	79.1 (56.1)	87.1 (57.7)	94.8 (98.9)	100.2 (242.9)
Close$^+$1	73.3 (46.8)	80.7 (49.6)	87.9 (72.3)	92.9 (95.0)
Close$^+$2	72.1 (45.8)	79.5 (48.8)	86.5 (68.3)	91.5 (84.4)
Close$^+$3	71.5 (42.8)	78.7 (45.4)	85.7 (58.0)	90.6 (68.8)
Close$^+$4	71.0 (40.4)	78.2 (43.2)	85.1 (51.0)	90.0 (56.2)
Close$^+$5	70.6 (37.9)	77.8 (40.1)	84.6 (47.2)	89.5 (49.9)
Close$^+$10	70.5 (33.6)	77.7 (34.1)	84.6 (37.3)	89.5 (38.4)
Close$^+$15	72.0 (30.2)	79.3 (30.4)	86.3 (30.8)	91.2 (30.6)
Close$^+$20	70.8 (25.2)	77.9 (25.4)	84.8 (26.0)	89.6 (26.2)
Close$^+$25	70.5 (22.0)	77.6 (22.0)	84.4 (22.5)	89.1 (22.9)
Close$^+$29	70.8 (19.4)	77.9 (19.3)	84.7 (19.7)	89.4 (19.7)

4.2 Regular Trading Hours

Figure 5 plots the cumulative impulse response θ_t starting with the first market open (with dotted lines representing ± 2 standard errors). Figure 5b shows the entire first trading day and Figure 5a zooms in on the first hour. The starting point of both plots is θ_t estimated from opening auction prices. Figure 5a shows rapid adjustment during the first hour of regular trading, especially in the first few minutes. Figure 5b shows continued price discovery throughout the entire day, albeit at a slower rate than immediately after the open. This pattern is consistent with the theoretical results of Back & Pedersen (1998), Collin-Dufresne & Fos (2012*b*). They show that if noise-trader volume is time-varying, informed agents choose to trade when it is high. Consequently, prices impound information at a faster rate during times of high liquidity. Table 2 presents estimates of $\theta_{\omega:t}$ over time intervals ($\theta_{\omega:t} \equiv \theta_t - \theta_\omega$). For example, the first column reports 7.15 which equals $\theta_{9:30} - \theta_{\text{Open}+0} = \theta_{\text{Open}+0\ :\ 9:30}$. It is the change in cumulative price reaction from the opening auction price to the 9:30-9:33 VWAP. The results in Table 2 show that price adjustment to news continues throughout the entire trading day in an economically and statistically significant way. The anomaly is the last column, which measure price reaction during the closing twenty minutes of trading. The non-monotonicity observed in Figure 5a is statistically significant though small. I conjecture this retracing of prices is due to a discontinuity in trading costs at 16:00. NASD Rule 2520 allows day traders to employ up to 400% leverage for intraday equity positions. Positions held overnight are subject to the usual 200% leverage restriction. This discontinuity at the regular market close greatly increases the cost of holding a position after 16:00. The non-monotonic CIRF near the close is consistent with a binding leverage constraint for at least some informed traders who take advantage of increased liquidity near the close of trading to "zero out" their portfolios. The is similar to the finding in Cushing & Madhavan (2000) that increased demand for immediacy near the close generates return reversals. Reversal in CIRF_t near the close is further evidence that only a few specialized investors process and trade on information. This provides some support for binding limits to arbitrage other than risk-aversion.

Figure 5: Cumulative Impulse Response

(a) Event Day (1st Hour Regular Trading)

(b) Event Day (Regular Trading)

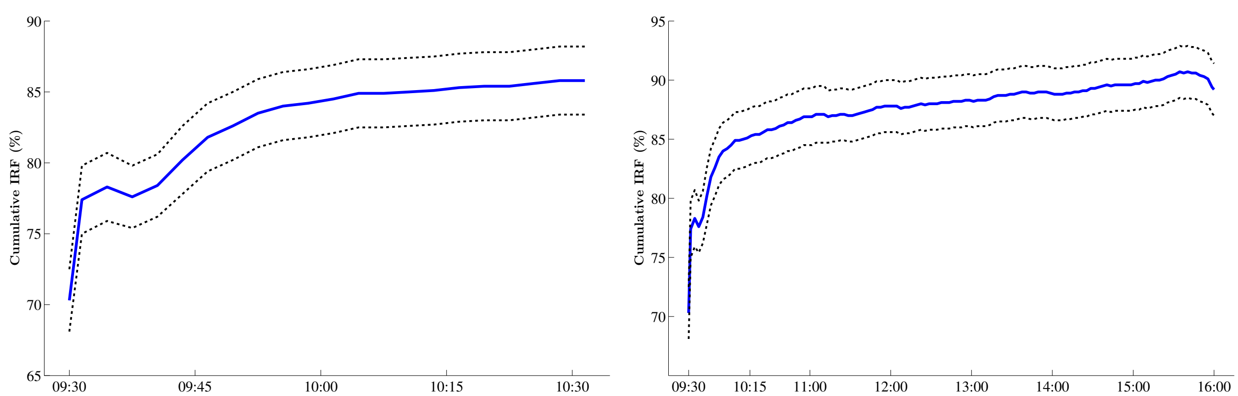


Table 2: Event Day Impulse Response

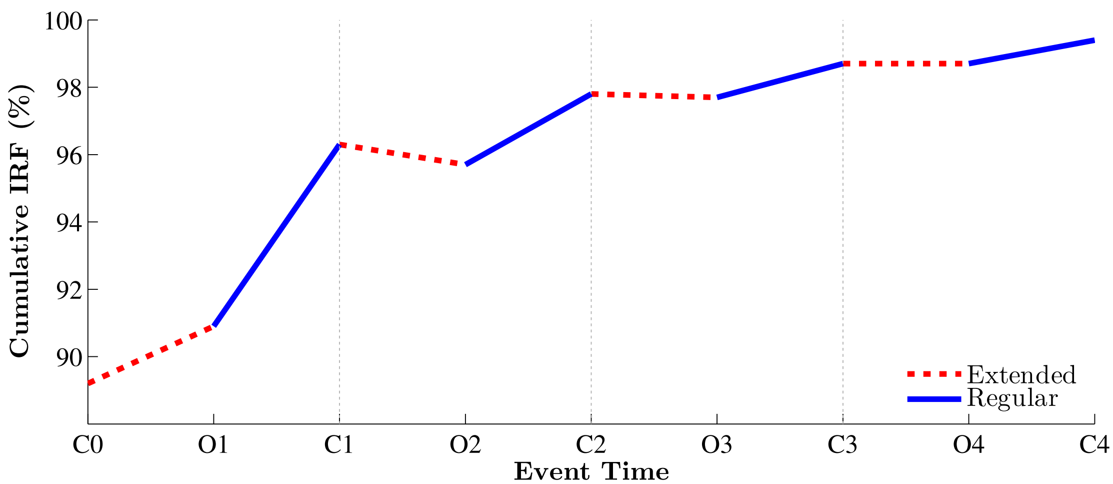
Each column displays the change in cumulative impulse response (in percent) measured over the given horizon. For example, the third column reports 1.27 for the period 10:00-10:30, meaning the cumulative impulse response increases by 1.27% during this time frame.

Standard errors are adjusted for heteroskedasticity and are clustered by firm and day of announcement. Absolute t statistics in parentheses.

	Open-9:30	9:30-10:00	10:00-10:30	10:30-11:30	11:30-1:00	1:00-2:30	2:30-3:40	3:40-Close
$\theta_{\omega:t}$	7.15***	6.90***	1.31***	1.22**	1.17**	1.15***	1.44***	−1.57***
	(17.1)	(7.9)	(3.3)	(2.7)	(2.9)	(3.3)	(4.8)	(7.7)
N	34269	34269	34269	34269	34269	34269	34269	34269
# Firms	3060	3060	3060	3060	3060	3060	3060	3060
# Year x Month	72	72	72	72	72	72	72	72

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 6: Post-Event Days 1-4



I decompose price discovery in the four days post-event into regular and extended hours trading sessions. Figure 6 plots θ_t starting at the close of the first post-event trading day (where Figure 5b ends). The dotted red segments correspond to price adjustment from market close to the next day open. Solid blue segments correspond to price reaction during regular trading hours. So each pair of red then blue segment represent a full day. Except for the first overnight session, price adjustment takes place during regular trading hours. This is seen from the essentially flat red-dotted segments and upward sloping blue-solid segments. Table 3 presents estimates of $\theta_{\omega:t}$ from Close-to-Open or Open-to-Close (as indicated) on event days $^{+}1$ to $^{+}4$. For example, the first column reports 1.68 which equals $\theta_{\text{Open}^{+}1} - \theta_{\text{Close}^{+}0}$. It is the change in cumulative impulse response from the closing auction on event day to the opening auction the next morning (the height of the first red dotted segment in Figure 6). The estimates confirm the pattern in Figure 6 is statistically significant. Each

Table 3: Trading vs Non-trading Hours Impulse Response

Each column measures the price adjustment from Close-to-Open or Open-to-Close (as indicated) on event days 1-4. For example, the fourth column reports 2.15 for the period 10:00-10:30, meaning the cumulative impulse response increases by 2.15% during trading hours on event day $^{+}2$.

Standard errors are adjusted for heteroskedasticity and are clustered by firm and month of announcement. Absolute t statistics in parentheses.

	$C^{+}0 - O^{+}1$	$O^{+}1 - C^{+}1$	$C^{+}1 - O^{+}2$	$O^{+}2 - C^{+}2$	$C^{+}2 - O^{+}3$	$O^{+}3 - C^{+}3$	$C^{+}3 - O^{+}4$	$O^{+}4 - C^{+}4$
$\theta_{\omega:t}$	1.68***	5.26***	−0.61	2.14***	−0.17	1.06*	−0.01	0.75
	(4.42)	(7.19)	(1.60)	(3.40)	(0.55)	(1.71)	(0.03)	(1.14)
N	34269	34269	34269	34269	34269	34269	34269	34269
# Firms	3060	3060	3060	3060	3060	3060	3060	3060
# EA Days	1412	1412	1412	1412	1412	1412	1412	1412

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

day, most if not all of the price adjustment occurs during regular trading. Except for the first overnight ($C^{+0} - O^{+1}$), the estimated adjustment outside trading hours is zero (or negative). In contrast, there is significant positive price reaction during regular trading hours each day. This pattern is further evidence that strategic informed investors wait to trade when liquidity is predictably higher.

4.3 Extended Hours, Trade by Trade

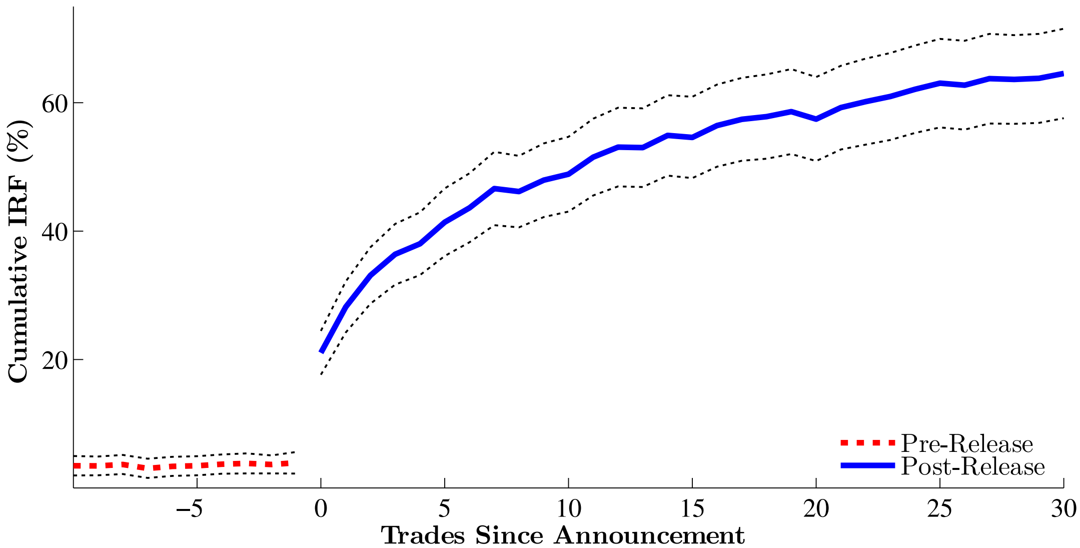
Gradual diffusion predicts that CIRF_t is approximately equal to the fraction of agents who are informed at time t. To quantitatively assess this prediction, I estimate CIRF_t trade by trade. In order to obtain a relatively homogenous sample, I restrict to announcements made after market close for the remainder of the paper[16]. Because both the earnings data from WSH and trade data from TAQ are timestamped only to the nearest second, I exclude trades that are timestamped within ±1 second of the earnings announcement. Results are essentially unchanged if I include all trades. I require at least thirty trades between the announcement and 18:30.

Figure 7 shows the estimated impulse response function around the announcement. Impulse response greater than zero at trade -1 is either due to information leaking after the close of regular trading but prior to the announcement or due to a potential bias in my timing data[17]. The most striking feature of Figure 7 is the small jump from trade -1 to trade 0. The CIRF jumps to only 21%. Taken literally, gradual diffusion would claim that at least 79% of agents are inattentive at the first trade but are not aware of their information disadvantage. The CIRF increases rapidly to 65% by the thirtieth trade (4.5 minutes post-announcement).

[16] Unless otherwise specified

[17] WSH records when the earnings press release comes over the newswires. Occasionally, a company posts the information on its website just prior to the newswire release.

Figure 7: Tick-by-Tick Cumulative IRF



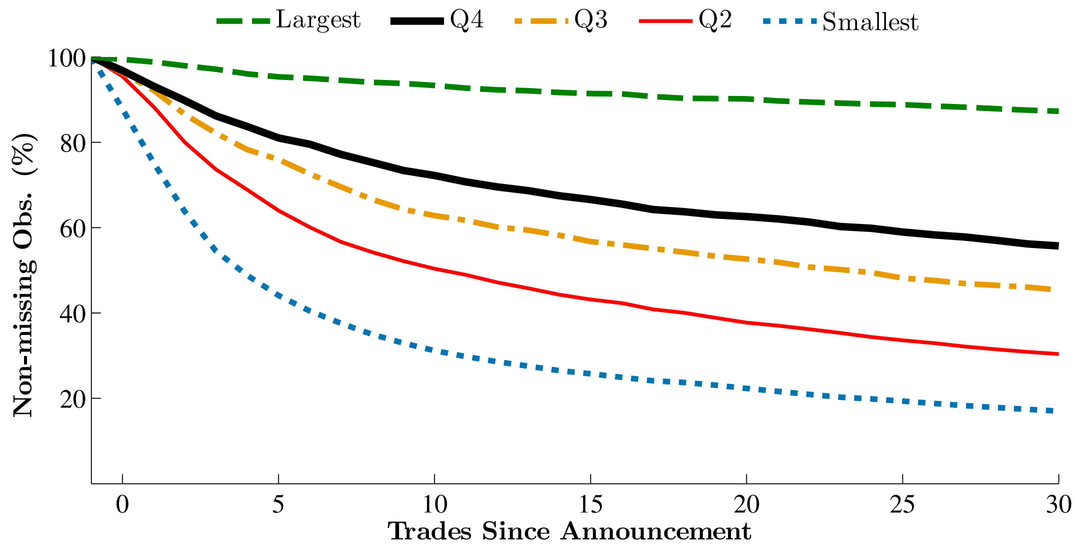
Trade -1 is the last trade prior to the earnings press release. Trade 0 is the first trade post announcement. Dotted black lines show two standard error bounds.

Inattention implies that of the 79% who initially ignored the news, half become informed by this time.

Figure 8 shows the trading survival functions by firm size. Trade -1 is the last trade prior to the earnings press release and trade 0 indicates the first trade after the announcement. The curves give the fraction of observations which have at least the indicated number of trades post-announcement. For instance, the highest curve shows that 90% of large firm announcements are accompanied by at least thirty after-hours trades after the earnings press release. Restricting to announcements that ex-post have thirty trades after the announcement may introduce selection bias since the probability of trading varies predictably in the cross section. I address this issue below.

I reestimate the cumulative impulse response at each horizon using all events with a non-missing price to address this concern. Figure 9a shows this function along with the estimates from Figure 7. The two functions essentially overlap, indicating that estimates using only observations with non-missing data capture the essential features of the impulse response. Figure 9b shows CIRF_t in the first hour of regular session trading separately for two groups: those observations with a valid thirtieth trade (solid blue) and those without (red dotted). Firms with non-missing after-hours prices have significantly higher CIRF at the open. Within one hour the difference is largely eliminated as the remaining firms catch up. This highlights the importance of extended hours trading for price discovery as well as

Figure 8: Tick-by-Tick Survival Functions by Firm Size



Trade -1 is the last trade prior to the earnings press release. Trade 0 is the first trade post-announcement.

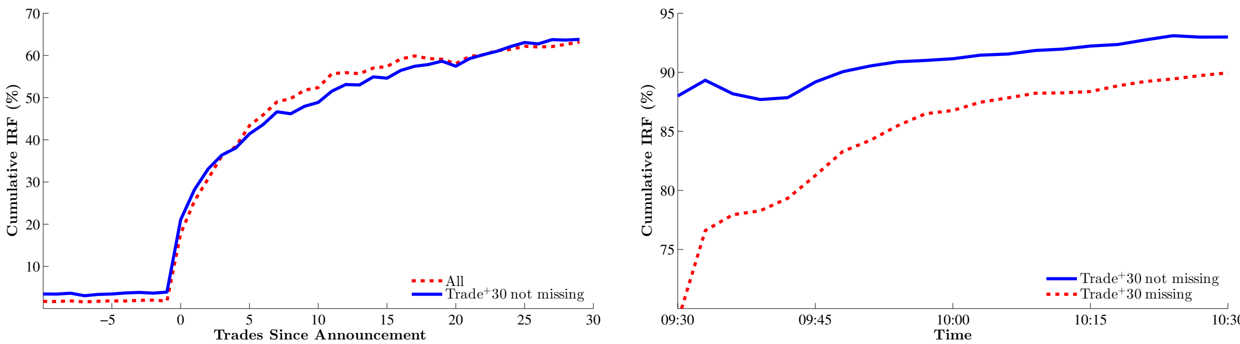
the phenomenon that jumps in liquidity are associated with increases in the *rate* of price discovery. This observation may be subject to selection bias since I sort firms on whether they have non-missing data. In Section 5 I sort firms on ex-ante observables which predict low after-hours trading. These sorts generate functions similar to Figure 9b.

4.3.1 Clock Time vs Trading Time

The slow diffusion hypothesis predicts that as time passes more investors switch from inattentive to informed and therefore price converges to true value. The strategic trading hypothesis predicts that the mere passage of time is irrelevant for beliefs; agents only learn through observing prices and quantities of shares actually traded. In other words, conditional on the "trading clock", the impulse response should be independent of the "real clock." To test this hypothesis, I sort firms based on the time elapsed from the earnings release to the thirtieth trade into two equal size groups (Fast and Slow), within each size quintile, within each calendar quarter (Year × Quarter). This method orthogonalizes speed relative to firm size and calendar time which is important since speed is strongly negatively correlated with firm size. Median elapsed time for Fast and Slow firms are 70 and 950 seconds, respectively. Table 4 shows this sort is orthogonal to other variables of interest, providing a clean measure of speed. This mitigates concerns of selection bias since I sort events on ex-post elapsed time.

Figure 9: Selection Bias

(a) Tick-by-Tick (b) Regular Trading



Trade -1 is the last trade prior to the earnings release. Trade 0 is the first trade post announcement. "All" is estimated from all events with a valid trade at the indicated horizon. "Trade^{+}30 not missing" is from Figure 7.

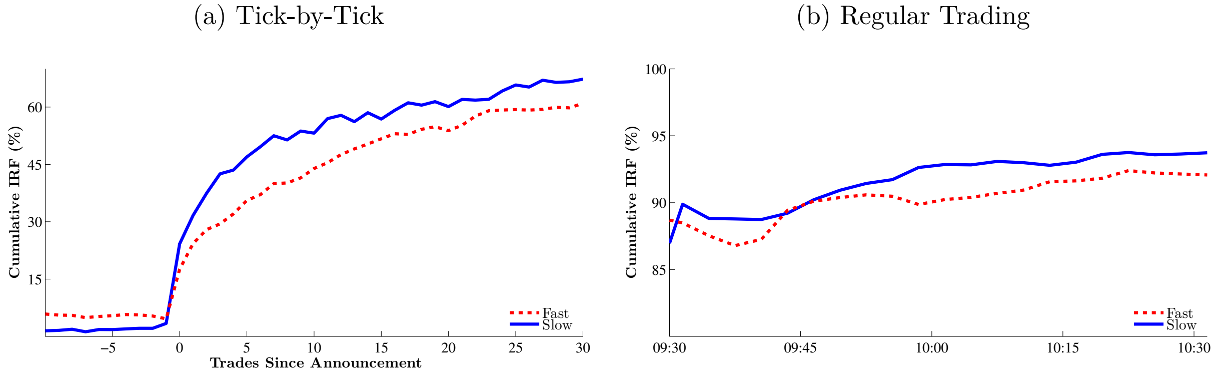
Table 4: Summary Statistics by Trading Speed

Mean, median, and standard deviation for indicated measure by trading speed.

	Mean		Median		σ	
	Fast	Slow	Fast	Slow	Fast	Slow
Time to 30th Trade (Seconds)	143.33	1156.22	70.00	951.00	194.13	917.33
Surprise	0.15	0.09	0.04	−0.00	0.82	0.82
$\text{CAR}_{\text{Close}+5}$	−0.25	−0.26	−0.11	−0.36	12.23	11.46
$\text{CAR}_{\text{Close}+20}$	−0.29	−0.22	−0.05	−0.09	15.76	14.59
Institutional Ownership (%)	0.79	0.79	0.81	0.84	0.18	0.19
Mean Analyst Recommendation	2.39	2.39	2.40	2.40	0.58	0.62
ME Quintile	3.26	3.25	3.00	3.00	1.39	1.39
BE/ME Quintile	2.02	2.15	2.00	2.00	1.10	1.15

Figure 10a shows $\text{CIRF}_{t,\text{Fast}}$ and $\text{CIRF}_{t,\text{Slow}}$ around the earnings announcement; Prices of Slow firms do adjust somewhat faster but the two functions are fairly similar. Table 5 provides formal tests of equality; the difference between the two functions is marginally significant at short horizons but decays after the 5th trade. Figure 10b shows the functions during the first thirty minutes of regular trading (~17 hours after the announcement); they are nearly identical. For slow diffusion to explain this pattern, speed of learning for the two types of firms must be drastically different for the first few minutes after an announcement (since elapsed time to thirtieth trade differs by an order of magnitude), but nearly identical for the next day (since $\text{CIRF}_{\text{Open}+0}$ is equal across types).

Figure 10: Cumulative IRF by Trading Speed



Trade -1 is the last trade prior to the earnings press release. Trade 0 is the first trade post announcement. Firms are sorted by duration from announcement to thirtieth trade, within year, quarter, and size quintile.

Table 5: Cumulative IRF by Trading Speed

$$\mathbb{E}\left[CAR_t \middle| S, \text{Speed}\right] = a + \theta_{\text{Fast}} \cdot \mathbb{E}\left[CAR_T \cdot \mathbf{1}_{\text{Fast}} \middle| S, \text{Speed}\right] + \theta_{\text{Slow}} \cdot \mathbb{E}\left[CAR_T \cdot \mathbf{1}_{\text{Slow}} \middle| S, \text{Speed}\right]$$

Each column shows the cumulative impulse response at the indicated trade (relative to earnings announcement). For example, the second column shows that at the thirtieth trade after the announcement, fast firms CIRF is θ=65.3% and slow firms have an additional 2.2% CIRF.

Standard errors are adjusted for heteroskedasticity. Absolute t statistics in parentheses.

	Trd_{-1}	Trd_{+0}	Trd_{+5}	Trd_{+10}	Trd_{+20}	Trd_{+30}
θ_{Fast}	4.54***	17.56***	35.59***	43.93***	53.84***	60.99***
	(3.67)	(7.86)	(10.43)	(11.13)	(12.07)	(12.38)
$\theta_{\text{Slow}} - \theta_{\text{Fast}}$	−1.11	6.64**	11.39**	9.26	6.27	6.30
	(0.64)	(1.96)	(2.17)	(1.61)	(0.97)	(0.91)
N	4157	4157	4157	4157	4157	4157

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 Cross-Section

Kyle & Obizhaeva (2011) develop a market microstructure invariance hypothesis based on the idea that time-series and cross-sectional variation in market processes essentially disappears when measured in "business time" instead of "calendar time." In Section 4.3 I showed that conditional on equal trading, CIRF_t is independent of the "real clock". In this section I provide further evidence little cross-sectional variation in speed of price discovery exists when measured in business time. I show that business time is well measured by the number of trades since announcement, not volume or turnover. Invariance in the cross-section is an equilibrium outcome. Indeed, in Kyle's model, CIRF_t depends only on the number of auctions and is independent of the amount of noise trading and information precision even though uninformed agents learn from order flow. This results obtains since in equilibrium informed trading volume is exactly proportional to noise trader volume.

5.1 Liquidity and Competition

Holden & Subrahmanyam (1992) extend the model of Kyle (1985) to include multiple informed traders. They find that increasing competition leads to faster price adjustment. In the perfectly competitive limit, all private information is incorporated into prices instantaneously. In a model of endogenous information acquisition, Kyle (1989) shows that the number of informed agents is increasing in the volume of noise trading. The endogenous correlation between liquidity and competition blurs interpretation of tests which assess the impact of each factor. I take an agnostic approach and do not attempt to seperately measure their effects.

I measure cross-sectional variation in price discovery using various measures of liquidity and informed competition. I use holdings data from quarterly 13F filings to sort firms into two groups each calendar quarter based on institutional ownership. If retail investors are subject to inattention then HighRetail should capture high behavioral bias as well as reduced informed competition. Since active traders are the clients of equity analysts, the number of analysts may be a good proxy for the level of competition among active investors. Further, more analyst coverage is thought to reduce information asymmetry through broader distribution of information. I define a event as having low coverage if only one analyst provided an earnings estimate for that quarter. As one measure of liquidity, I sort events based on trading volume on the firms' *previous* earnings announcement day. Event day volume provides a better measure of news-induced noise trading than average volume over non-news days. I calculate a modified version of the price impact measure from (Amihud 2002) as $\mathbb{E}\left[\frac{|r_t|}{Turnover_t}\right]$, using daily data in the prior calendar year. As with volume, I sort firms into two bins each quarter. Bid/ask spread is a commonly used measure of asymmetric information based on the theory in Glosten & Milgrom (1985). I calculate average daily closing effective spread, $\mathbb{E}\left[\frac{1}{2}\frac{ask-bid}{ask+bid}\right]$, for each firm in the prior calendar year and again sort announcements each quarter into two bins. Sorting by quarter should remove any potential trends in the proxies which may be related to other unobserved firm characteristics. Previous studies have concluded that post-earnings announcement drift is a phenomenon concentrated in small firms (Bernard & Thomas 1990, Hou 2007, Chordia et al. 2009). This is rationalized by the observations that small firms have a higher percentage of (potentially biased) retail investors and higher trading costs which may deter informed arbitrageurs from trading aggressively. In addition to the aforementioned liquidity and competition measures, I use market equity quintiles since firm size is often used as a catch all for market frictions.

Table 6 shows summary stats by measure. In addition to the variables used to construct the liquidity measures, I include two measures using options data. From OptionMetrics' volatility surface file, I extract the implied volatility of a 1-month at-the-money call option as of the last market close prior to the earnings announcement. In the last row I show the fraction of announcements with no traded options (missing implied volatility). Implied volatility is, as expected, decreasing in firm size. In the last row I show the fraction of announcements with no traded options (missing implied volatility). Nearly half of announcements in the smallest quintile do not have traded options, whereas nearly all other firms do. The first five columns show that firm size is correlated with each of the other variables, but most strongly with Amihud measure and bid/ask spread[18]. From the rest of the table we see that the liquidity and competition measures are not uncorrelated, but the competition measures generate variation along a different dimension than the liquidity measures.

I estimate the following specifications using the same IV method described in Section 4.1:

$$\begin{aligned} \theta_t &= \kappa_t + b_{1,t} \cdot \mathbf{1}_{\text{ME}=1} + b_{2,t} \cdot \mathbf{1}_{\text{ME}=2} + b_{3,t} \cdot \mathbf{1}_{\text{ME}=3} + b_{4,t} \cdot \mathbf{1}_{\text{ME}=4} \qquad (3) \\ \theta_t &= \kappa_t + b_t \cdot \mathbf{1}_{\text{Liq/Comp Measure}} \end{aligned}$$

with a corresponding first-stage estimation in which the earnings surprise, $\mathbf{S}$, is interacted with the liquidity/competition proxy (see Appendix C). Table 7 shows the estimated CIRF_t from Open_{+0} to Close_{+0} with standard errors in parentheses. κ_t is CIRF_t for the more liquid or competitive group and $b_{j,t}$ is the difference in CIRF between the illiquid and liquid subsamples. $b_{j,t}$ for each measure shows a large spread at the open which substantially decays over the trading day, particularly during the first hour. For firms in size quintile 5, the opening price reflects essentially 100% of the earnings news. This is consistent with the findings of Lee & Gerlach (2011), who study stocks in the DJIA and measure speed of adjustment using excess volatility. For smaller firms there is rapid, but incomplete convergence within the first thirty minutes of trading. The spread at open and rapid decay is further evidence that jumps in liquidity are accompanies by an increase in the rate of price discovery. Measured in calendar time, there is large cross-sectional variation in the speed of price discovery.

Table 8 shows θ_t trade-by-trade. The spreads in CIRF_t observed in Table 7 are eliminated or even reversed when measured in tick time. Small and illiquid firms adjust slightly faster to news when measured per trade though the difference is not significant. When measuring prices tick-by-tick after hours, the discreteness of order quantities is relevant. Glosten & Milgrom (1985) solve a discrete version of the model in Kyle (1985). They have informed

[18] This can also be seen from the first row. The spread in log market equity by measure is strongest for Amihud and bid/ask.

Table 6: Summary Statistics by Liquidity/Competition Measure

Each column gives mean and median (in parentheses) values of the listed variables for firms in the indicated market equity quintile.

	Market Equity					HighRetail		LowCoverage		LowVolume		HighAmihud		HighSpread	
	Small	2	3	4	Big	0	1	0	1	0	1	0	1	0	1
ln(MktEquity)	5.64	6.70	7.52	8.36	9.89	7.22	6.72	7.41	6.33	7.27	6.78	7.67	6.22	7.83	6.09
	(5.72)	(6.72)	(7.53)	(8.33)	(9.61)	(7.13)	(6.34)	(7.25)	(6.23)	(7.16)	(6.49)	(7.53)	(6.12)	(7.66)	(6.03)
Institutional Ownership	0.65	0.79	0.82	0.82	0.74	0.90	0.54	0.78	0.70	0.81	0.70	0.82	0.67	0.82	0.68
	(0.69)	(0.85)	(0.87)	(0.85)	(0.76)	(0.90)	(0.58)	(0.83)	(0.75)	(0.86)	(0.74)	(0.86)	(0.70)	(0.86)	(0.71)
# Analysts=1	0.57	0.37	0.26	0.17	0.04	0.30	0.44	0.00	1.00	0.25	0.50	0.22	0.53	0.22	0.52
	(1.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)	(1.00)
Event Day Turnover (%)	2.09	2.95	3.20	2.87	3.00	3.12	2.11	3.27	1.69	3.89	1.26	3.53	1.69	3.10	2.24
	(0.96)	(1.76)	(2.18)	(1.99)	(2.15)	(2.04)	(1.07)	(2.14)	(0.86)	(2.79)	(0.81)	(2.47)	(0.90)	(2.13)	(1.05)
Amihud Illiquidity	9.45	3.54	2.52	1.62	1.47	2.58	8.16	3.37	7.44	2.50	6.73	1.48	8.86	1.70	8.35
	(4.29)	(2.34)	(1.71)	(1.46)	(1.34)	(1.79)	(3.38)	(1.78)	(3.29)	(1.63)	(3.31)	(1.46)	(4.20)	(1.55)	(3.98)
Bid/Ask Spread (%)	0.40	0.18	0.13	0.10	0.07	0.15	0.33	0.17	0.30	0.16	0.26	0.11	0.35	0.10	0.36
	(0.26)	(0.15)	(0.11)	(0.09)	(0.07)	(0.12)	(0.21)	(0.12)	(0.19)	(0.12)	(0.18)	(0.10)	(0.23)	(0.10)	(0.24)
Implied Volatility	0.62	0.56	0.50	0.44	0.43	0.51	0.55	0.53	0.53	0.55	0.49	0.49	0.58	0.47	0.62
	(0.58)	(0.52)	(0.47)	(0.41)	(0.40)	(0.48)	(0.52)	(0.49)	(0.49)	(0.51)	(0.45)	(0.46)	(0.55)	(0.44)	(0.58)
No Traded Options	0.40	0.10	0.02	0.00	0.00	0.09	0.28	0.10	0.30	0.07	0.26	0.03	0.33	0.03	0.32
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

agents and noise traders who can each trade only a single unit and a monopolist market maker who chooses a bid/ask spread. In their model, more adverse selection (higher ratio of informed to noise traders) leads to wider bid/ask spreads and more price adjustment *conditional* on a trade occurring. They show that if adverse selection is too large no trade occurs, as in the market for "lemons" (Akerlof 1970). Indeed, Figure 8 shows that the after hours trading small stocks tends to halt soon after the announcement. So *per trade*, small stocks will adjust to news more quickly, but since large firms trade much more, *per unit time* large stocks respond faster.

Table 7: CIRF_t by Liquidity/Competition Measure

Each column measures the cumulative impulse response (in %) by proxy at the indicated horizon. $\theta_{\mathrm{t}} = \kappa_t + b_t' \mathbf{1}_{\mathrm{Liq/Comp}}$. Standard errors are adjusted for heteroskedasticity. Absolute t statistics in parentheses.

	Open^{+}0	9:30	10:00	10:30	12:30	Close^{+}0
κ	96.2	94.2	95.3	93.9	99.8	102.8
	(10.4)	(10.7)	(11.0)	(11.3)	(12.0)	(12.0)
b_1	−32.0***	−23.0**	−14.3	−11.4	−12.5	−14.8*
	(3.3)	(2.5)	(1.6)	(1.3)	(1.4)	(1.7)
b_2	−19.1*	−10.1	−4.7	0.1	−2.3	−3.7
	(1.9)	(1.1)	(0.5)	(0.0)	(0.2)	(0.4)
b_3	−12.6	−4.1	3.7	5.7	−1.7	−3.0
	(1.2)	(0.4)	(0.4)	(0.6)	(0.2)	(0.3)
b_4	−10.9	−5.0	3.0	4.7	−1.8	−3.1
	(0.9)	(0.4)	(0.3)	(0.4)	(0.2)	(0.3)
κ	80.0	86.0	93.9	96.3	98.3	97.7
	(28.7)	(29.9)	(33.4)	(34.4)	(37.0)	(38.8)
$b_{\mathrm{HighRetail}}$	−13.5***	−12.9***	−11.4***	−12.5***	−9.9**	−7.0*
	(3.5)	(3.2)	(2.9)	(3.1)	(2.5)	(1.9)
κ	81.6	86.5	92.7	94.3	96.7	97.9
	(30.3)	(31.4)	(34.6)	(35.6)	(37.9)	(40.0)
$b_{\mathrm{LowCoverage}}$	−18.9***	−15.6***	−10.1**	−9.5**	−7.6*	−8.3**
	(4.9)	(3.9)	(2.5)	(2.4)	(2.0)	(2.2)
κ	84.9	89.9	95.3	97.3	99.5	99.4
	(29.0)	(30.2)	(33.6)	(34.5)	(36.8)	(38.6)
$b_{\mathrm{LowVolume}}$	−22.3***	−18.8***	−11.4***	−11.1***	−9.4**	−7.6*
	(5.5)	(4.4)	(2.7)	(2.6)	(2.3)	(1.9)
κ	83.1	87.8	95.0	97.3	100.7	100.9
	(27.4)	(28.2)	(31.1)	(31.9)	(34.2)	(36.0)
$b_{\mathrm{HighAmihud}}$	−16.7***	−13.9***	−11.8***	−12.5***	−12.7***	−11.5***
	(4.2)	(3.4)	(2.9)	(3.1)	(3.3)	(3.1)
κ	84.0	89.3	95.1	97.4	100.5	101.6
	(28.0)	(28.8)	(31.5)	(32.4)	(34.7)	(36.8)
$b_{\mathrm{HighSpread}}$	−16.7***	−15.1***	−10.8***	−11.6***	−11.3***	−11.5***
	(4.3)	(3.7)	(2.7)	(2.9)	(2.9)	(3.1)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: CIRF_t by Liquidity/Competition Measure

Each column measures the cumulative impulse response (in %) by proxy at the indicated horizon. $\theta_{\mathrm{t}} = \kappa_t + b_t' \mathbf{1}_{\mathrm{Liq/Comp}}$. The last column gives estimates (in %) from $\mathrm{Prob}\left[\mathrm{Trd}_{30}\ \mathrm{Missing}\right] = \kappa + b' \mathbf{1}_{\mathrm{Liq/Comp}}$. Standard errors are adjusted for heteroskedasticity. Absolute t statistics in parentheses.

	Trd_{-1}	Trd_{+0}	Trd_{+10}	Trd_{+20}	Trd_{+30}	$\mathrm{Trd}_{+30} = \varnothing$
κ	7.1	21.5	42.5	51.7	58.7	12.7
	(3.0)	(5.4)	(6.8)	(7.0)	(7.2)	(13.5)
b_1	−2.2	2.5	15.1	13.0	14.6	70.3***
	(0.6)	(0.4)	(1.5)	(1.1)	(1.2)	(61.4)
b_2	−7.0**	−2.9	6.9	11.5	10.3	56.9***
	(2.1)	(0.5)	(0.7)	(1.0)	(0.8)	(44.5)
b_3	−4.0	−2.4	6.9	3.8	2.6	41.9***
	(1.3)	(0.5)	(0.7)	(0.4)	(0.2)	(27.9)
b_4	−1.7	3.2	8.5	8.3	4.7	31.6***
	(0.6)	(0.5)	(0.7)	(0.6)	(0.3)	(19.9)
κ	3.4	20.6	47.3	57.2	64.2	58.4
	(3.4)	(9.7)	(13.3)	(14.2)	(14.9)	(95.2)
$b_{\mathrm{HighRetail}}$	0.3	0.6	3.7	0.5	0.7	6.5***
	(0.2)	(0.2)	(0.6)	(0.1)	(0.1)	(6.7)
κ	4.2	20.5	47.3	56.5	63.3	51.6
	(4.4)	(11.0)	(14.9)	(15.9)	(16.6)	(88.2)
$b_{\mathrm{LowCoverage}}$	−2.4	2.8	10.3	7.1	8.0	29.8***
	(1.3)	(0.6)	(1.2)	(0.8)	(0.8)	(33.4)
κ	3.8	20.4	46.1	55.5	62.9	49.5
	(4.0)	(11.2)	(15.3)	(16.2)	(17.2)	(82.3)
$b_{\mathrm{LowVolume}}$	−0.4	1.3	10.6	8.2	5.7	32.7***
	(0.1)	(0.2)	(0.9)	(0.7)	(0.4)	(37.6)
κ	4.2	18.5	44.9	54.3	62.0	44.4
	(4.4)	(9.7)	(14.1)	(14.9)	(15.8)	(70.5)
$b_{\mathrm{HighAmihud}}$	−1.1	9.8**	14.7*	12.5	11.0	39.9***
	(0.5)	(2.1)	(1.9)	(1.5)	(1.2)	(47.8)
κ	3.9	19.8	45.8	54.7	61.9	46.7
	(3.8)	(9.8)	(13.5)	(14.3)	(15.0)	(72.1)
$b_{\mathrm{HighSpread}}$	−0.1	3.2	8.1	7.5	7.8	32.0***
	(0.0)	(0.8)	(1.2)	(1.0)	(1.0)	(36.4)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: CIRF_t Conditional on $\text{Trd}_{+30} \neq \varnothing$

Each column measures the cumulative impulse response (in %) by proxy at the indicated horizon. $\theta_{\text{t}} = \kappa_t + b_t' \mathbf{1}_{\text{Liq/Comp}}$. Sample restricted to announcements with $\text{Trd}_{+30} \neq \varnothing$. Standard errors are adjusted for heteroskedasticity. Absolute t statistics in parentheses.

	Open^{+}0	9:30	10:00	10:30	12:30	Close^{+}0
κ	94.4	94.9	99.1	98.1	100.4	102.2
	(10.9)	(11.1)	(11.2)	(11.8)	(12.2)	(11.5)
b_1	−7.4	−6.2	−6.5	−4.2	1.8	1.3
	(0.6)	(0.5)	(0.5)	(0.4)	(0.2)	(0.1)
b_2	−1.8	−2.0	−6.6	−2.2	−1.7	1.4
	(0.1)	(0.2)	(0.6)	(0.2)	(0.2)	(0.1)
b_3	−9.2	−8.4	−10.4	−8.8	−11.9	−12.6
	(0.7)	(0.7)	(0.9)	(0.7)	(1.0)	(1.0)
b_4	−5.6	−3.3	−2.0	2.3	1.4	0.3
	(0.4)	(0.2)	(0.1)	(0.1)	(0.1)	(0.0)
κ	88.7	90.1	92.4	95.2	97.2	98.0
	(19.7)	(20.6)	(22.7)	(23.1)	(24.3)	(25.1)
$\text{b}_{\text{HighRetail}}$	−2.7	−3.4	−3.3	−7.1	−3.3	−0.5
	(0.3)	(0.4)	(0.5)	(1.0)	(0.5)	(0.1)
κ	88.6	90.4	91.6	93.3	96.5	98.6
	(22.6)	(23.2)	(25.5)	(25.9)	(27.3)	(28.3)
$\text{b}_{\text{LowCoverage}}$	−2.2	−4.6	0.2	−0.4	1.5	2.0
	(0.2)	(0.5)	(0.0)	(0.0)	(0.2)	(0.2)
κ	88.9	90.2	91.4	93.7	97.6	99.7
	(22.8)	(23.5)	(25.6)	(26.1)	(27.5)	(28.5)
$\text{b}_{\text{LowVolume}}$	−5.1	−5.3	4.0	0.8	−1.6	0.8
	(0.4)	(0.5)	(0.3)	(0.1)	(0.2)	(0.1)
κ	86.1	87.5	89.9	92.7	97.2	98.4
	(21.5)	(22.0)	(24.0)	(24.4)	(25.5)	(26.4)
$\text{b}_{\text{HighAmihud}}$	7.9	7.4	6.6	1.8	−2.2	0.4
	(0.8)	(0.8)	(0.8)	(0.2)	(0.3)	(0.1)
κ	88.2	89.9	92.2	94.1	97.3	98.8
	(20.9)	(21.5)	(23.4)	(23.7)	(24.8)	(25.5)
$\text{b}_{\text{HighSpread}}$	0.2	−0.8	−1.1	−1.9	−1.4	0.3
	(0.0)	(0.1)	(0.2)	(0.3)	(0.2)	(0.0)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

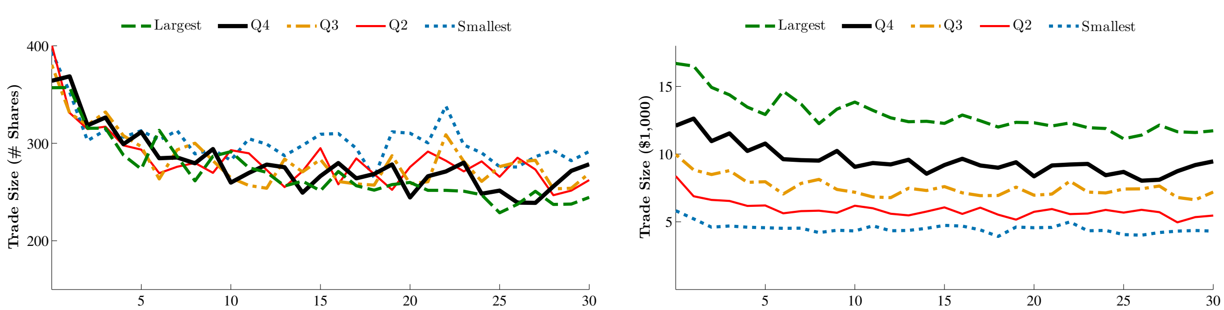
So what is it that causes the spreads in $\text{CIRF}_{\text{Open}+0}$ observed in Table 7? The last column of Table 8 shows the probability that thirty extended hours trades post-event will not occur by proxy; the low liquidity and low competition subgroups have significantly less after-hours trading volume. Figure 9b showed that missing Trd_{30} leads to lower CIRF at Open_{+0}. Table 9 repeats the estimation from Table 7 except now I restrict to the sample of announcements with Trd_{30} not missing. Compared to Table 7, the spreads are reduced by more than half and none are statistically significant. This result highlights the importance of after-hours trading to price discovery. “Calendar time” variations in the speed of price discovery disappear after conditioning on equal “business time”.

5.2 Trades, Shares, Dollars, Turnover?

The previous section shows that CIRF_t is approximately invariant to firm size, and measure of liquidity and asymmetric information. I now explore whether trade time is best captured by number of transactions or other measures of trading activity. Figure 11 plots trading volume tick-by-tick averaged within market equity quintile[19]. Figure 11a shows volume measured in number of shares per transaction. As during regular trading hours, the distribution of trades is concentrated at “even lots”, i.e. 100, 200, 300, etc (not shown). There is very little difference in trade size across market equity quintiles. Volume drops for the first 5-10 trades then stabilizes. Figure 11b plots mean dollar volume and shows a similar pattern with one substantial difference. Since firm size is correlated with share price, dollar volume shows an increasing relation with firm size whereas share volume seems to be independent of size. The largest firms trade 2-3 times the dollar volume of the smallest firms but are 50 times as large as measured by median market capitalization. The fifth row gives the standard deviation of $\widehat{\mathbb{E}}\left[CAR_{\text{Close}+5}\middle| S\right]$ by size quintile, a measure of the typical size of the fundamental value shock. The smallest firms experience shocks that are typically two times the size as those experience by large firms. Measuring “drifts” in return space would show that small firms have much higher predictable returns even though Table 8 shows that CIRF_t is approximately equal for firms of all sizes.

Figure 11: Trading Volume

(a) Shares (b) Dollars



Trade 0 is the first trade post-announcement. Values are 5% trimmed means by market equity quintile.

[19] I trim the data at the 5th and 95th percentiles to reduce the impact of outliers. I use trimmed means instead of medians since discreteness in trade size makes median values very sensitive to sampling error.

Table 10: Trading Statistics by Firm Size

The first three rows are median values by market equity quintile for various measures over the first thirty trades post-announcement. Rows four and five are standard deviations of predicted returns at indicated horizons. The final row is the ratio of row four and row three.

	Small	2	3	4	Big
Market Equity ($ MM)	390.58	899.19	1887.18	4472.59	17159.84
Trading Volume ($ K)	163.11	238.89	296.58	368.25	426.28
Volume/Equity (b.p.)	4.47	2.64	1.60	0.79	0.22
$std\left[\widehat{\mathbb{E}}\left(CAR_{\text{Trd+30}} \mid S\right)\right]$ (%)	3.48	2.63	2.11	1.61	1.39
$std\left[\widehat{\mathbb{E}}\left(CAR_{\text{Close+5}} \mid S\right)\right]$ (%)	4.83	4.01	3.44	3.05	2.38
Elasticity $\partial log(P)/\partial log(Q)$	77.81	99.57	131.83	203.73	639.55

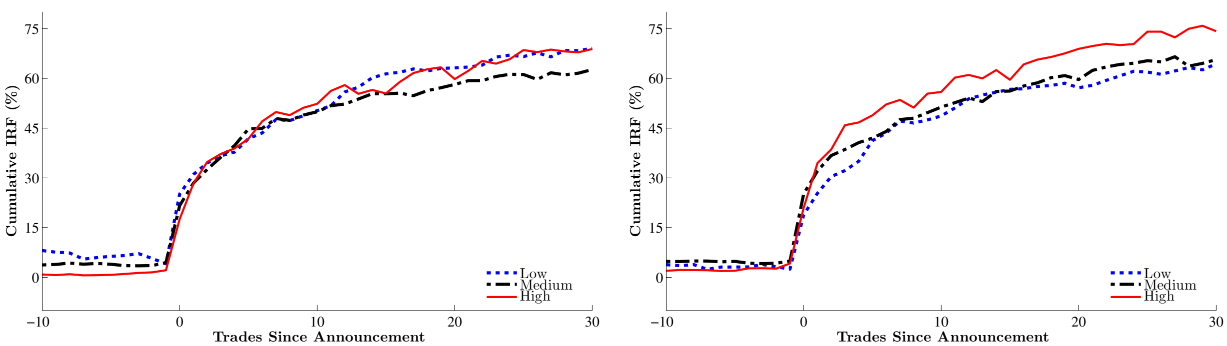
Table 10 gives median trading statistics by market equity quintile (for the thirty trades post-announcement). Though dollar volume for the largest firms is approximately 2.5 times that for small firms, turnover for small firms is larger by a factor of 20. Given the approximate invariance in trade-by-trade CIRF_t shown in Table 8, it is obvious that turnover is not an approriate measure of trade time. Two remaining candidates are share volume and dollar volume. To test these, I sort announcements into three groups by total share volume and dollar volume over the thirty trades. For the groups “Low”, “Medium” and “High”, median share volumes are 6500, 11000, 21500 shares and median dollar volumes are $147K, $300K, $668K, respectively. If trading volume is important for price discovery, these sorts should generate differences in CIRF_t when t is measured in transactions since announcement. Figure 12 again shows invariance of CIRF_t to these cross-sectional dimensions, suggesting that price discovery occurs as a function of trades, not volume. This relates to the findings of Jones et al. (1994) who show that “it is the occurrence of transactions per se, and not their size, that generates volatility.” Using expected prices instead of variances, I show that transactions, and not their size, generate price discovery. Though Kyle & Obizhaeva (2011) highlight the distinction between order flow and the actual passage of business time, the approximate invariance in $\text{CIRF}_{\text{Trd+30}}$ suggests a tight link between the two. Jones et al. (1994) emphasize the point that “both the frequency and size of trades are endogenously determined and yet the size of trades has no information content beyond that contained in the number of transactions.”

One possible theoretical approach to address this phenomenon is modelling markets as “quote driven”, as in Glosten & Milgrom (1985). In Kyle (1985) and related papers, markets are “order driven”. Traders place quantity orders and market clearing prices are determined. Informed agents optimize along an intensive margin, choosing the size of trades, but not number of trades. This methodology is attractive primarily due to the analytical tractability of such models. In high-frequency, real asset markets are more quote driven. Liquidity providers submit limit orders for a given price and quantity and traders then choose whether or not to accept these offers. In such a market, informed agents optimize on an extensive margin, chooosing not the size of, but rather, the number of trades. Quote driven markets are characterized by discreteness in quantities and price changes which adds substantial difficulty in solving the models.

Figure 12: CIRF$_t$ by Trading Volume

(a) Share Volume

(b) Dollar Volume



Trade 0 is the first trade post-announcement

5.3 Risk-aversion or Adverse Selection?

An important difference between diffusion and strategic trading is the reason for downward sloping uninformed asset demand. In strategic trading models, uninformed agents recognize their information disadvantage and learning from order flow generates inelastic demand. Since uninformed agents do not learn from prices in the diffusion model, risk-aversion drives the slope of their demand curve. In this section I explore whether risk-aversion is sufficient to generate the large after-hours price changes with small trading volume seen in **??**, **??** and **??**. To quantitatively address this issue I estimate a stylized model of liquidity provision without adverse selection.

Assume a single risky asset in zero net supply and normalize the gross risk-free rate to unity. At date 0 a random liquidity shock arrives an at date 1 a liquidating dividend is paid. From the perspective of uninformed agents in the diffusion model, price changes result from random noise trader shocks so this setup provides a suitable laboratory to study price impact. Further, asset returns are normally distributed and a continuum of liquidity providers have power utility over wealth, $U_0 = \mathbb{E}_0\left[\frac{W_1^{1-\gamma}}{1-\gamma}\right]$. Standard portfolio optimization yields $\omega = \frac{\mu}{(\gamma-1)\sigma^2}$ where μ and σ^2 are the conditional mean and variance of returns given the beliefs of liquidity providers and ω is the fraction of wealth invested in the risky asset. By definition, the amount of liquidity provision equals $W_0\omega$ where W_0 is the total initial wealth of the uninformed. Let L be the total dollar amount of liquidity provision. Then $L = W_0\omega$ which implies $W_0 = \frac{L(\gamma-1)\sigma^2}{\mu}$. From this expression, it is clear that any L, μ σ^2 can be rationalized with either high risk-aversion γ or low initial wealth W_0.

I set L equal to the median total dollar volume from announcement to the thirtieth trade (see Table 10). I choose $\mu = \theta_{\text{Trd+30}}\widehat{\mathbb{E}}\left[\text{CAR}_{\text{T}}\right]$. This gives the "typical" predicted price change from announcement to the thirtieth trade, which underestimates actual price changes. I assume uninformed agents believe prices converge to fundamental value within one trading day (by Close$^+$0) since sufficient risk-bearing capacity to fully absorb the liquidity shock arrives during the regular trading day. Thus, I set σ^2 equal to the estimated variance of returns from Trd$^+$30 to Close$^+$0. Following Bansal et al. (2012) I choose $\gamma = 10$. Table 11 gives estimates of W_0 by market equity quintile. It ranges from \$330K for the smallest stocks to \$2.2M for the largest. These are implausibly small values for the total wealth of liquidity providers after hours, suggesting that risk-aversion alone cannot generate the observed price impact of trading. Another way to see this is by looking at implied Sharpe ratios. These are calculated from the perspective of the uninformed agents and range from 2.6 (annualized) for large firms to 6.6 for small firms. If liquidity provision is perceived as so profitable it begs the question of why there isn't more entry into market making after hours. The adverse selection explanation for large price impact does not suffer from these issues. As shown in Kyle (1985), price impact tends to infinity as noise trader volume goes to zero, even with risk-neutral perfectly competitive market makers.

Table 11: Liquidity Provision

Implied wealth of uninformed agents and perceived Sharpe ratio by market equity quintile.

	1	2	3	4	5
Implied W_0 (\$1000)	\$331	\$624	\$1,024	\$1,378	\$2,194
Sharpe Ratio (μ/σ^2)	0.42	0.33	0.25	0.23	0.16
Sharpe Ratio (annualized)	6.61	5.17	3.94	3.61	2.61

6 Conclusion

Using precise event timing and high-frequency trade data, I measure the impulse response of stock prices to public news with high resolution. I employ a structural approach which nests competing models of price discovery. This allows me to contrast their qualitative and quantitative ability to match the estimated price process. I find that asset prices incorporate information through trading, though significant volume is not required when markets are illiquid. Price discovery occurs at a faster rate when markets are more liquid. The cumulative impulse response function seems to have kinks whenever there are jumps in liquidity. I find that extended hours trading contributes significantly to price discovery. The efficiency of the opening price after an announcement is strongly related to the amount of extended hours trading prior to the open. I show that little cross-sectional variation in the speed of price discovery exists when measured in trade time, consistent with the invariance hypothesis of Kyle & Obizhaeva (2011). Prices impound information through the occurrence of transactions, independent of trade size. The characteristics of price discovery after a public news release are consistent with models of price formation in response to private information. Understanding why public information induces belief heterogeneity is an important open question.

References

Akerlof, G. A. (1970), 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *The Quarterly Journal of Economics* pp. 488–500.

Amihud, Y. (2002), 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets* **5**(1), 31–56.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Vega, C. (2007), 'Real-time price discovery in global stock, bond and foreign exchange markets', *Journal of International Economics* **73**(2), 251–277.

Angrist, J. D. (1998), 'Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants', *Econometrica* **66**(2), 249–288.

Back, K. & Pedersen, H. (1998), 'Long-lived information and intraday patterns', *Journal of Financial Markets* **1**(3), 385–402.

Ball, R., Sadka, G. & Sadka, R. (2009), 'Aggregate earnings and asset prices', *Journal of Accounting Research* **47**(5), 1097–1133.

Banerjee, S. (2011), 'Learning from Prices and the Dispersion in Beliefs', *Review of Financial Studies* **24**(9), 3025–3068.

Bansal, R., Kiku, D., Shaliastovich, I. & Yaron, A. (2012), Volatility, the Macroeconomy and Asset Prices, Technical report, National Bureau of Economic Research.

Barberis, N., Shleifer, A. & Vishny, R. (1998), 'A model of investor sentiment', *Journal of Financial Economics* **49**(3), 307–343.

Barclay, M. J. & Hendershott, T. (2004), 'Liquidity Externalities and Adverse Selection: Evidence from Trading after Hours', *The Journal of Finance* **59**(2), 681–710.

Barclay, M. J. & Hendershott, T. (2008), 'A comparison of trading and non-trading mechanisms for price discovery', *Journal of Empirical Finance* **15**(5), 839–849.

Beechey, M. J. & Wright, J. H. (2009), 'The high-frequency impact of news on long-term yields and forward rates: Is it real?', *Journal of Monetary Economics* **56**(4), 535–544.

Bernard, V. & Thomas, J. (1990), 'Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings', *Journal of Accounting and Economics* **13**(4), 305–340.

Bessembinder, H. (1999), 'Trade Execution Costs on Nasdaq and the NYSE: A Post-Reform Comparison', *Journal of Financial and Quantitative Analysis* **34**(3).

Brandt, M. W. & Kavajecz, K. A. (2004), 'Price Discovery in the US Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve', *The Journal of Finance* **59**(6), 2623–2654.

Brogaard, J., Hendershott, T. J. & Riordan, R. (2012), High Frequency Trading and Price Discovery.

Carlin, B. I., Longstaff, F. A. & Matoba, K. (2012), Disagreement and Asset Prices.

Chordia, T., Goyal, A., Sadka, G., Sadka, R. & Shivakumar, L. (2009), 'Liquidity and the Post-Earnings-Announcement Drift', *Financial Analysts Journal* pp. 18–32.

Chordia, T. & Shivakumar, L. (2006), 'Earnings and price momentum', *Journal of Financial Economics* **80**(3), 627–656.

Collin-Dufresne, P. & Fos, V. (2012*a*), Do prices reveal the presence of informed trading?

Collin-Dufresne, P. & Fos, V. (2012*b*), Insider Trading, Stochastic Liquidity and Equilibrium Prices.

Cushing, D. & Madhavan, A. (2000), 'Stock Returns and Trading at the Close', *Journal of Financial Markets* **3**(1), 45–67.

Daniel, K., Hirshleifer, D. & Subrahmanyam, A. (1998), 'Investor Psychology and Security Market Under-and Overreactions', *The Journal of Finance* **53**(6), 1839–1885.

DellaVigna, S. & Pollet, J. (2009), 'Investor inattention and Friday earnings announcements', *The Journal of Finance* **64**(2), 709–749.

Evans, M. D. & Lyons, R. K. (2008), 'How is macro news transmitted to exchange rates?', *Journal of Financial Economics* **88**(1), 26–50.

Evans, M. & Lyons, R. K. (2002), 'Order Flow and Exchange Rate Dynamics', *Journal of Political Economy* **110**(1), 170–180.

Fama, E. F. (1998), 'Market efficiency, long-term returns, and behavioral finance', *Journal of Financial Economics* **49**, 283–306.

Fleming, M. J. & Remolona, E. M. (1999), 'Price formation and liquidity in the US Treasury market: The response to public information', *The journal of Finance* **54**(5), 1901–1915.

Foster, F. D. & Viswanathan, S. (1996), 'Strategic Trading When Agents Forecast the Forecasts of Others', *The Journal of Finance* **51**(4), 1437–1478.

Frazzini, A. (2006), 'The Disposition Effect and Underreaction to News', *The Journal of Finance* **61**(4), 2017–2046.

Gabaix, X. (2012), A Sparsity-Based Model of Bounded Rationality, Applied to Basic Consumer and Equilibrium Theory.

Glosten, L. R. & Milgrom, P. R. (1985), 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of financial economics* **14**(1), 71–100.

Green, J. R. . (1973), Information, Efficiency, and Equilibrium. Harvard Institute of Economic Research Discussion Paper No. 284.

Green, T. C. (2004), 'Economic News and the Impact of Trading on Bond Prices', *The Journal of Finance* **59**(3), 1201–1234.

Harris, M. & Raviv, A. (1993), 'Differences of opinion make a horse race', *Review of Financial Studies* **6**(3), 473–506.

Hayek, F. (1945), 'The Use of Knowledge in Society', *The American Economic Review* **35**(4), 519–530.

Heiner, R. A. (1983), 'The Origin of Predictable Behavior', *The American Economic Review* **73**(4), 560–595.

Hirshleifer, D. & Teoh, S. H. (2003), 'Limited attention, information disclosure, and financial reporting', *Journal of Accounting and Economics* **36**(1), 337–386.

Holden, C. & Subrahmanyam, A. (1992), 'Long-Lived Private Information and Imperfect Competition', *The Journal of Finance* **47**(1), 247–270.

Hong, H. & Stein, J. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *The Journal of Finance* **54**(6), 2143–2184.

Hong, H. & Stein, J. (2007), ‘Disagreement and the stock market’, *The Journal of Economic Perspectives* **21**(2), 109–128.

Hou, K. (2007), ‘Industry Information Diffusion and the Lead-lag Effect in Stock Returns’, *Review of Financial Studies* **20**(4), 1113–1138.

Huang, R. D. & Stoll, H. R. (1996), ‘Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE’, *Journal of Financial economics* **41**(3), 313–357.

Jones, C. M., Kaul, G. & Lipson, M. L. (1994), ‘Transactions, volume, and volatility’, *Review of Financial Studies* **7**(4), 631–651.

Kelley, E. K. & Tetlock, P. C. (2012), Why Do Investors Trade?

Kim, D. & Kim, M. (2003), ‘A multifactor explanation of post-earnings announcement drift’, *Journal of Financial and Quantitative Analysis* **38**(2), 383–398.

Kim, O. & Verrecchia, R. (1994), ‘Market liquidity and volume around earnings announcements’, *Journal of Accounting and Economics* **17**(1), 41–67.

Kothari, S. (2001), ‘Capital markets research in accounting’, *Journal of Accounting and Economics* **31**(1), 105–231.

Kyle, A. S. (1985), ‘Continous Auctions and Insider Trading’, *Econometrica* **53**, 1315–1335.

Kyle, A. S. (1989), ‘Informed speculation with imperfect competition’, *The Review of Economic Studies* **56**(3), 317–355.

Kyle, A. S. & Obizhaeva, A. (2011), Market Microstructure Invariants: Theory and Implications of Calibration.

Lee, J. & Gerlach, J. (2011), Is the Opening Price Efficient? Evidence from After-Hours Earnings Announcements.

Sadka, R. (2006), ‘Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk’, *Journal of Financial Economics* **80**(2), 309–349.

Shin, H. S. (2006), ‘Disclosure Risk and Price Drift’, *Journal of Accounting Research* **44**(2), 351–379.

Stein, J. C. (2009), ‘Presidential address: Sophisticated investors and market efficiency’, *The Journal of Finance* **64**(4), 1517–1548.

Appendix

A Models of Price Discovery

Table 12: Models of Price Formation

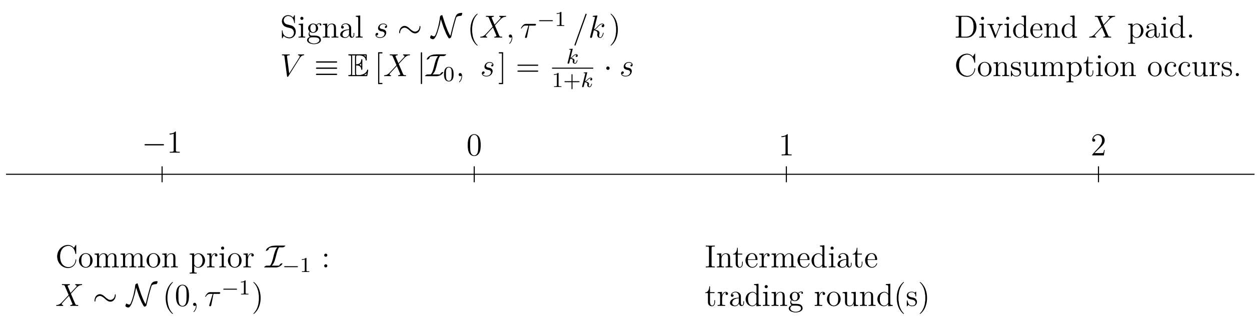
		Uninformed	
		Agree to Disagree	Learn from Trading
Informed	Competitive	Hong & Stein (1999)	Grossman (1977)
	Strategic		Kyle (1985)

Taxonomy of models of price discovery.

Table 12 shows a taxonomy of models of price discovery. The rows are distinguished by whether "informed" agents are price-takers as in the noisy rational expectations literature or behave strategically (endogenize price impact) as in Bayesian Nash equilibrium models. The columns differ according to the learning behavior of uninformed agents. Papers such as Grossman (1977) and Kyle (1985) assume that uninformed agents learn optimally using market data (prices and quantities). Others assume that uninformed "agree to disagree"; that is, they enter the marketplace with beliefs and do not update based on observations of trading activity. Slow diffusion/inattention models such as Hong & Stein (1999), DellaVigna & Pollet (2009) typically assume that informed agents are competitive price-takers and uninformed agents don't learn from market data. Kyle (1985) assumes the polar opposite; informed traders are strategic and uninformed update their beliefs using prices and quantities according to Bayes' rule. Running a "horse race" of the models is difficult since the physical environments differ in the original papers. In this section I solve the models in a common setting to generate predictions that depend only on the behavioral assumptions employed. Additionally, the common context allows me to isolate the impact of each dimension of Table 12 on the equilibrium properties of price and trading volume.

The world has four dates, denoted $t = -1 \dots 2$; Figure 13 shows a timeline of the environment. There is a single risky asset in zero net supply which pays a random dividend, X,

Figure 13: Model Timeline



at t_2. At t_{-1}, all agents have a common prior belief, $\mathcal{I}_{-1}:\ X \sim \mathcal{N}\left(0, \tau^{-1}\right)$. At t_0 there is a publicly observable signal $s \sim \mathcal{N}\left(X, \tau^{-1}/k\right)$[20]. The Bayesian posterior belief conditional on the signal is $V \equiv \mathbb{E}\left[X \,|\mathcal{I}_0,\ s\right] = \frac{k}{1+k} \cdot s$ and $\mathrm{Var}\left(X \,|\mathcal{I}_{-1},\ s\right) = \left(\tau\left(1+k\right)\right)^{-1}$. The gross riskless rate is normalized to unity. Agents may trade freely on all dates and price at each date is determined by a Walrasian auctioneer. I define the cumulative impulse response function as

$$\theta_t \equiv \mathrm{CIRF}_t \ \equiv \ \mathbb{E}\left(\frac{P_t - P_{-1}}{V - P_{-1}} \middle| \mathcal{I}_0,\ s \right)$$

which captures the fraction of the shock which is reflected in price.

A.1 Canonical Diffusion/Inattention (Competitive, Agree to Disagree)

There is a continuum of investors indexed on $i \in [0, A]$. They have identical preferences given by $U_{i,t} = -\mathbb{E}_{i,t}\left[exp\left(-\rho W_{i,2}\right)\right]$; this is exponential utility with coefficient of absolute risk aversion ρ[21]. At t_0, a fraction μ_0 of investors observe the signal s and update their beliefs regarding the payoff X. The remaining $1-\mu_0$ fraction do not observe the signal and, *crucially*, are unaware of its existence. They do not learn about X from prices or order quantities; their posterior beliefs remain fixed at the prior $X \sim \mathcal{N}\left(0, \tau^{-1}\right)$ no matter what happens in the asset market. "Uninformed agents" view the price change as resulting from a permanent liquidity shock/noise trader shock[22]. At t_1 information may diffuse to more agents so that $\mu_1 \geq \mu_0$. Define $\mu_{-1} \equiv 0$ since at t_{-1} no agents have observed s. Similarly,

[20] Equivalently, $s = X + \varepsilon,\ \varepsilon \sim \mathcal{N}\left(X, \tau^{-1}/k\right)$

[21] Both DellaVigna & Pollet (2009) and (Stein 2009) use quadratic utility over next period wealth. With normally distributed payoffs the two preferences yield identical solutions.

[22] For simplicity, I assume that either (1) uninformed employ buy and hold strategies or (2) uninformed mistakenly believe liquidity shocks (noise trader demand) have zero variance.

$\mu_2 \equiv 1$ since all uncertainty is resolved at t_2.

A.1.1 "Buy and Hold" Informed

Hong & Stein (1999) assume informed agents "formulate their asset demands based" on a buy and hold strategy; they do not intend on trading at any future dates. In actuality they do rebalance so this assumption results in time-inconsistency of informed beliefs/behavior. The authors state it is merely a simplifying device and do not believe it affects any "important qualitative conclusions." For any μ_0, μ_1, the model has a simple solution given by: $P_{-1} = 0,\ P_0 = \theta_0 V,\ P_1 = \theta_1 V,\ P_2 = X$, where $\theta_t \equiv \frac{(1+k)\mu_t}{1+k\mu_t}$[23]. As seen in Figure 14a the solution exactly fits the diffusion intuition; CIRF_t exactly tracks μ_t, the fraction of the population which is informed. Dynamic inconsistency occurs as follows: at t_1 the t_0 informed sell some of their asset holding to the newly informed agents. However, they do not anticipate doing this at t_0 even though it occurs with certainty. I now relax the buy and hold assumption and allow informed agents to fully dynamically optimize.

A.1.2 Optimizing Informed

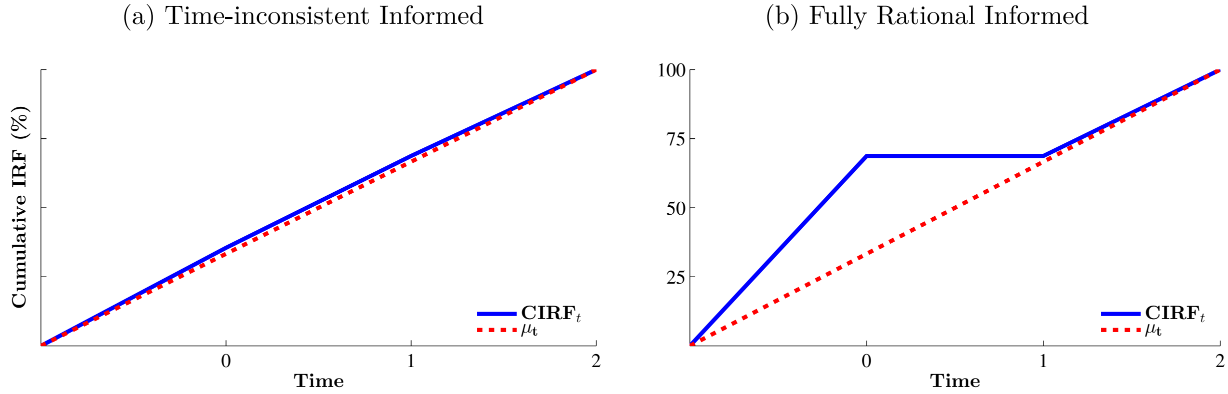
When informed agents are allowed to dynamically optimize, their decisions become time-consistent. This requires them to form beliefs about P_1, the price at the intermediate trading date. Fortunately, the model has a simple analytic solution in this setup: $P_{-1} = 0,\ P_0 = P_1 = \theta_1 V,\ P_2 = X$ where $\theta_1 \equiv \frac{(1+k)\mu_1}{1+k\mu_1}$. The key difference arising from dynamic consistency is that $P_0 = P_1$; price is unchanged from t_0 to t_1. Why doesn't the diffusion intuition hold in this model? Why is there no price drift at t_1? From the perspective of a t_0 informed agent, P_1 (the price at t_1) is a non-random variable; he knows it will be exactly $\theta_1 V = \frac{(1+k)\mu_1}{1+k\mu_1} V$. Since the informed agents act competitively, the equilibrium price at t_0 must be $\theta_1 V$ discounted at the gross risk-free rate which is here normalized to unity. Hence, there is no price drift at t_1 even though average beliefs converge towards the truth. At t_1 there will be significant trading volume; the t_0 informed will offload some of their position to the newly informed agents[24]. The buy and hold assumption, far from being innocuous, contains the full substance of the diffusion mechanism.

With sufficient restrictions on position size the t_0 informed will be constrained, leading to lower impulse response and hence, drift at t_1. I explore this possibility quantitatively in **??**. As a preview, the total trading volume after hours is so low making it implausible that active informed agents are constrained by position limits or other limits to arbitrage. This example

[23] $\mu_t \leq \theta_t \equiv \frac{(1+k)\mu_t}{1+k\mu_t} \leq 1$. Closeness of θ_t to either μ_t or 1 depends on k, the signal to noise ratio of the signal s. $\lim_{k\to 0} \theta_t = \mu_t$ and $\lim_{k\to\infty} \theta_t = 1$.

[24] At each time t, there is perfect risk-sharing among the informed agents.

Figure 14: Inattention/Gradual Diffusion



highlights an important point about inattention models with atomistic agents. Price drift is a risk-premium from the perspective of informed agents. When there is no uncertainty resolved at a given date there can be no price drift at that date even if "information diffusion" occurs.

Figure 14a shows both CIRF_t and μ_t based on a calibration of $\mu_0 = 1/3$, $\mu_2 = 2/3$ and $k = 0.1$[25]; this matches the linear diffusion of Hong & Stein (1999). Figure 14a illustrates that diffusion does produce smooth price drift when informed agents have time-inconsistent behavior. Figure 14b shows that when informed agents dynamically optimize, early informed agents push prices to reflect learning by others which has yet to occur.

A.2 Imperfect Competition (Strategic, Agree to Disagree)

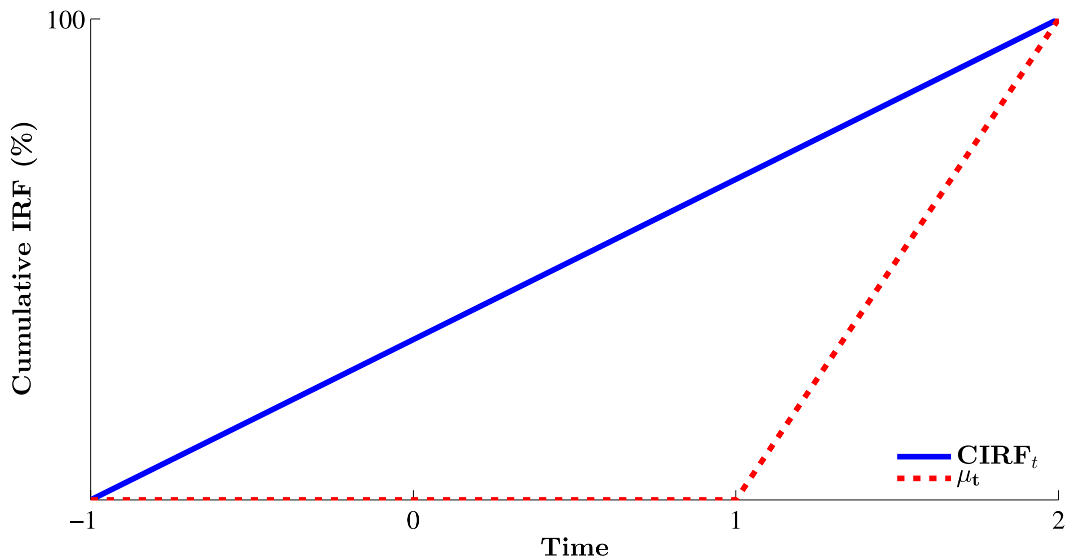
The previous models featured "perfect competition"; all agents are price-takers since they are each of zero measure. I relax this assumption by adding a risk-neutral "monopolist" informed agent. A fully rational risk-neutral agent M observes the signal s and fully endogenizes the impact of his trades on equilibrium prices. M is only allowed to place market orders [26]. He faces a "residual supply curve" which is just the aggregate demand of the A price-taking agents[27]. Since they are risk-averse, this supply curve will be upward sloping. For simplicity, I set $\mu_0 = \mu_1 = 0$ so that the atomistic agents are all "inattentive"; they never learn until the final payoff is revealed. Since inattentive agents do not learn, the supply curve will be

[25] I choose $\mu_1 = 0.5$ for legibility of the graph. I calibrate k by observing that $\frac{\mathbb{V}[X|\mathcal{I}_0,\ s]}{\mathbb{V}[X|\mathcal{I}_0]} = \frac{1}{1+k} \approx 0.9$ based on the average changes in implied volatility around an earnings announcement

[26] With risk-neutrality his optimal demand schedule takes the unpleasant form $Q = 0 \iff P = V$ and $|Q| = \infty$ *otherwise*. Brogaard et al. (2012) show that informed trades by high-frequency traders are market orders.

[27] Transformed to make it a supply curve.

Figure 15: Imperfect Competition



constant over t=-1..1 (at t_2 there is no uncertainty so we have the horizontal supply curve $P_2 = X$). This model features *no change* in liquidity conditions resulting from the release of signal s.

Equilibrium prices are given by: $P_{-1} = 0$, $P_0 = \frac{V}{3}$, $P_1 = \frac{2V}{3}$, $P_2 = X$. Figure 15 shows the cumulative impulse response function as well as the beliefs of the A price-taking inattentive agents. Notice that at t_1 the CIRF increases toward 100% even though *no* learning takes place at this date. Prices converge because the market is not infinitely liquid; the monopolist M trades based on his knowledge of s and since the A agents don't have infinite risk-bearing capacity, prices move. M does not trade an infinite quantity at t_0 because he recognizes the market is not infinitely liquid due to risk-aversion of A agents; M endogenizes the impact of his trades on equilibrium prices. Price impact creates a convex cost akin to risk-aversion even though M is risk-neutral[28]. The t_{-1} expected total trading volume on dates $t = 0 \ldots 2$ is proportional to ${}^{A}/_{\rho}$. This ratio represents the risk-bearing capacity of the A inattentive agents. Price discovery with little volume can occur only if inattentive agents are very risk-averse (high ρ) or have very little wealth (low A).

[28] M's objective function is quadratic in quantity, similar to a mean-variance investor.

A.3 Private Information (Strategic, Learn from Trading)

I modify the model of Section A.2 by changing the learning behavior of A agents. As before, I assume they don't update their beliefs based on signal s (the don't observe s) . However, I now do allow them to learn from prices and quantities. For analytical simplicity I assume they are risk-neutral, though the results are qualitatively unchanged with risk-aversion. In addition, I introduce "noise-traders" who have stochastic, perfectly inelastic asset demand given by $\ell_{-1} \equiv 0$ and $\ell_t \sim \mathcal{N}(0, \sigma^2)$ for $t \geq 0$. This is essentially the model of Kyle (1985) except that I specify a noisy signal whereas Kyle allows agent M to observe payoff X without error. With risk-neutrality this does not affect the essential character of the solution.

Let q_t be the number of shares purchased by M at time t. The restriction that M may only place market orders means that his demand is of the form $\forall\ t \geq 0: \quad q_t = f(s,\ P_{-1}, \ldots, P_{t-1},\ q_{-1}, \ldots, q_{t-1})$. An equilibrium is defined by an aggregate "supply" function, P_t, for the A agents at each date t and $M's$ demand functions $Q = \{q_0, \ldots, q_2\}$ such that the following conditions hold:

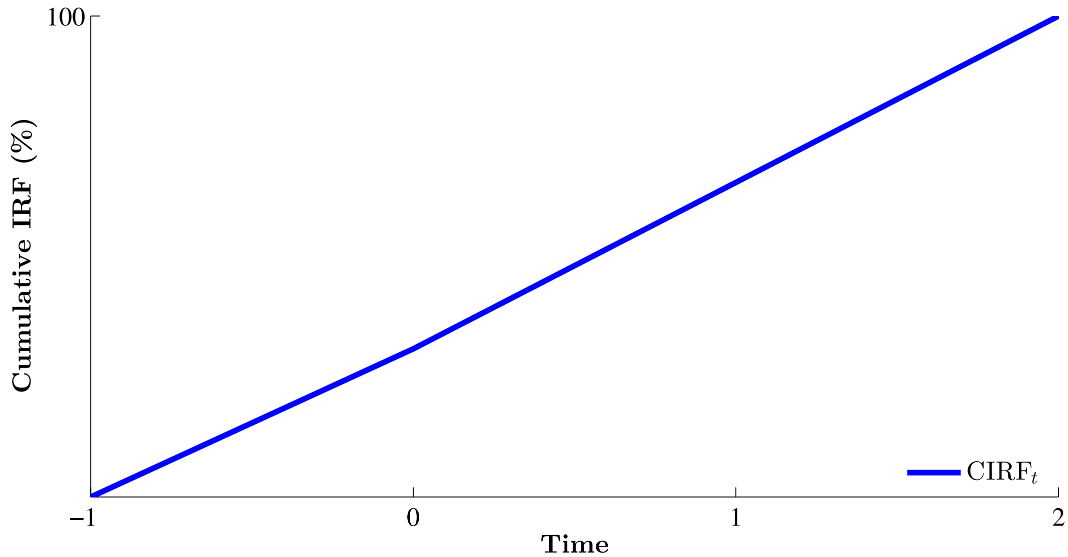
1. Market efficiency: Since the A agents are risk-neutral and competitive it holds that price equals their expected value of X.
   $P_t = \mathbb{E}[X \,| q_{-1} + \ell_{-1}, \ldots,\ q_t + \ell_t] = \mathbb{E}[V \,| q_{-1} + \ell_{-1}, \ldots,\ q_t + \ell_t]$. Expectations are taken conditional on the history of order imbalances (net order flow)[29].
2. Sequential rationality:
   $\forall\ t\ Q_t = \{q_t, \ldots, q_2\} = argmax\ \mathbb{E}[W_2 | \, s,\ P_{-1}, \ldots, P_{t-1},\ q_{-1}, \ldots, q_{t-1}]$ where W_2 is $M's$ terminal wealth. Conditional on any history, Q_t must be the optimal strategy for M in the remaining continuation game.

Appendix B describes the solution methodology; the equilibrium cannot be analytically determined though some properties can be. As seen in Figure 16, price discovery occurs almost linearly with $\text{CIRF}_{-1} = 0,\ \text{CIRF}_0 \approx 0.31,\ \text{CIRF}_1 \approx 0.65,\ \text{CIRF}_2 = 1$. Because of risk-neutrality, the impulse response does not depend on any parameters. Volume, however, does depend crucially on the amount of noise-trading. As $\sigma^2 \to 0$ total trading volume also tends to 0; λ_0, λ_1, the slopes of the uninformed supply curve tend to ∞. When uninformed learn from market data and there is very little noise trading, prices incorporate information with very little volume. Still, price discovery does require *actual* trading.

Holden & Subrahmanyam (1992) extend Kyle (1985) show that that imperfect competition (vs monopsonist) among informed agents increases the speed of price discovery; the CIRF will be more concave than in Figure 16. The result arises from the negative externality imposed by informed traders on each other; I recognize others are trading and price is con-

[29] The second equality comes from Law of Iterated Expectations: $\mathbb{E}[X \,| \{q_0 + z_0\}] = \mathbb{E}[\mathbb{E}(X \,| \{q_0 + z_0\},\ s) | \, \{q_0 + z_0\}] = \mathbb{E}[V \,| \{q_0 + z_0\}]$

Figure 16: Learning From Order Flow



verging towards V even if I do not trade. This increases my cost of delay and hence I trade more aggressively earlier. With endogenous costly information acquisition, the equilibrium number of informed agents is increasing in the volume of noise trading (Kyle 1989).

Collin-Dufresne & Fos (2012*b*) allow for time-varying noise trader volume (σ_2) and show informed agents choose to trade when there is more noise. Consequently, prices impound information at a faster rate during times of high liquidity. The CIRF should have "kinks" whenever there is a jump in liquidity (for example at the boundary between regular and extended hours trading). Using the trade record of Schedule 13D filers (SEC designated "insiders"), Collin-Dufresne & Fos (2012*a*) show that informed trading is in fact more aggressive during times of higher liquidity.

B Solution to Private Information Model

The backwards induction solution is as follows:

1. Conjecture linear aggregate supply functions of the form $P_t = P_{t-1} + \lambda_t [q_t + z_t]$ with boundary conditions $P_{-1} = 0$ and $P_2 = X$.
2. Solve $M's$ expected wealth maximization problem backwards from t=2 given the form of P_t (for arbitrary λ_1 and λ_2):

$$W_2 = q_0 [X - P_0] + q_1 [X - P_1] \tag{4}$$

$$q_1 = \frac{V - P_0}{2\lambda_1} \tag{5}$$

$$q_0 = \frac{V(\lambda_0 - 2\lambda_1)}{\lambda_0(\lambda_0 - 4\lambda_1)} \tag{6}$$

with associated second order conditions: $\lambda_0, \lambda_1 > 0$ and $\lambda_0 < 2\lambda_1$.

3. Calculate $\mathbb{E}_t[V] \equiv \mathbb{E}[V | q_{-1} + \ell_{-1}, \ldots,\ q_t + \ell_t]$ for t=0,1 as a function of λ_0 and λ_1 $\mathbb{E}_{-1}[V] = 0$. Let $\Sigma_t = \mathrm{Var}[V | q_{-1} + \ell_{-1}, \ldots,\ q_t + \ell_t]$ with $\Sigma_{-1} = \frac{k}{1+k}\tau^{-1}$.
   Let $j_0 = \frac{(\lambda_0 - 2\lambda_1)}{\lambda_0(\lambda_0 - 4\lambda_1)}$ and $j_1 = \frac{1}{2\lambda_1}$. $q_0 + z_0 \sim \mathcal{N}(j_0 V, \sigma^2)$.

$$P_0 = \mathbb{E}_0[V] = \frac{j_0 k}{(1+k)\sigma^2\tau + j_0^2 k} \cdot (q_0 + \ell_0) \tag{7}$$

$$P_1 = \mathbb{E}_1[V] = \mathbb{E}_0[V] + \frac{2k\lambda_1}{k + 4((1+k) + \sigma^2\tau + j_0^2 k)\lambda_1^2} \cdot (q_1 + \ell_1) \tag{8}$$

4. Solve numerically for λ_0 and λ_1 from $\mathbb{E}[V | q_{-1} + \ell_{-1}, \ldots,\ q_t + \ell_t] = 0$ for t=0,1
5. Substitute λ_0 and λ_1 into $\mathrm{CIRF}_t \equiv \mathbb{E}\left(\frac{P_t - 0}{V - 0}\middle| \mathcal{I}_{-1},\ V\right)$. Since $\mathbb{E}(\ell_t | \mathcal{I}_{-1},\ V) = 0$, $\mathrm{CIRF}_0 = j_0\lambda_0$ and $\mathrm{CIRF}_1 = \frac{1 + j_0\lambda_0}{2}$

C Identification

Recall the structural equations:

$$\forall\ t \geq -1,\ CAR_t^{i,q} = \theta_t^{i,q} \cdot \Delta V_{i,q} + \varepsilon_{i,q,t}$$

$$\theta_t^{i,q} = G_t(X_{i,q}) + \xi_{i,q,t},\ \mathbb{E}(\xi_{i,q,t}) = 0$$

$$\forall\ i,\ \forall\ t \geq T,\ \theta_t^{i,q} = 1$$

$$\forall t \geq -1,\ 0 = \mathrm{Cov}(z(S), \varepsilon_t) = 0,\ \mathrm{Cov}(\Delta V_{i,q}, \varepsilon_t) = 0$$

C.1 Pooled Estimation

Let $G_t(X_{i,q}) = \kappa_t$ so $\theta_t = \kappa_t + \xi_{i,q,t}$. For notational simplicity, let $r_t \equiv CAR_t^{i,q}$ and $r_T \equiv CAR_T^{i,q}$ and drop firm subscripts i, q. I derive the limiting values from 2SLS with a constant in the second stage using S as the first stage instrument:

$$\begin{aligned} z' &\equiv \begin{pmatrix} 1 & S \end{pmatrix} \\ x' &\equiv \begin{pmatrix} 1 & r_T \end{pmatrix} \\ y_t &\equiv r_t \end{aligned}$$

$$plim\ \widehat{\beta}_t = \mathbb{E}\left[zx'\right]^{-1}\mathbb{E}[zy]$$

$$
\begin{aligned}
&= \begin{pmatrix} 1 & \mathbb{E}[r_T] \\ \mathbb{E}[S] & \mathbb{E}[Sr_T] \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}[r_t] \\ \mathbb{E}[Sr_t] \end{pmatrix} \\
&= \frac{1}{\mathbb{E}[S\Delta V] - \mathbb{E}[S]\mathbb{E}[\Delta V]} \begin{pmatrix} \mathbb{E}[S\Delta V] & -\mathbb{E}[\Delta V] \\ -\mathbb{E}[S] & 1 \end{pmatrix} \begin{pmatrix} \mathbb{E}[\theta_t \Delta V] \\ \mathbb{E}[\theta_t S\Delta V] \end{pmatrix} \\
& \frac{1}{\mathbb{E}[S\Delta V] - \mathbb{E}[S]\mathbb{E}[\Delta V]} \begin{pmatrix} \mathbb{E}[SV]\mathbb{E}[\theta_t \Delta V] - \mathbb{E}[V]\mathbb{E}[\theta_t S\Delta V] \\ \mathbb{E}[\theta_t S\Delta V] - \mathbb{E}[S]\mathbb{E}[\theta_t \Delta V] \end{pmatrix}
\end{aligned}
$$

$$
\text{Let } \widehat{\theta}_t \equiv \widehat{\beta}_t^{(2)}
$$

$$
plim\ \widehat{\theta}_t = \frac{\mathbb{E}[\theta_t S\Delta V] - \mathbb{E}[S]\mathbb{E}[\theta_t \Delta V]}{\mathbb{E}[S\Delta V] - \mathbb{E}[S]\mathbb{E}[\Delta V]} = \frac{cov(S, \theta_t \Delta V)}{cov(S, \Delta V)}
$$

If θ_t is uncorrelated with ΔV and $S\Delta V$ then $\widehat{\theta}_t \underset{p}{\rightarrow} \mathbb{E}[\theta_t] = \kappa_t$ and the intercept converges to $\mathbb{E}[\varepsilon_t] - \theta_t \mathbb{E}[\varepsilon_T]$. Replacing the instrument S with any function $z(S)$ produces consistent estimates as long as it is correlated with V and θ is uncorrelated with $z(S)V$. Consider the case where $cov(\theta, SV) < 0$ which is true if firms that tend to have large value shocks also have slower price discovery. Suppose $\mathbb{E}[V],\ \mathbb{E}[\theta V] = 0$ for simplicity:

$$
\begin{aligned}
plim\ \widehat{\theta}_t &= \frac{\mathbb{E}[\theta_t SV] - \mathbb{E}[S]\mathbb{E}[\theta_t V]}{\mathbb{E}[SV] - \mathbb{E}[S]\mathbb{E}[V]} \\
&= \frac{\mathbb{E}[\theta_t]\mathbb{E}[SV] + \overbrace{cov(\theta_t, SV)}^{<0}}{\mathbb{E}[SV]} < \mathbb{E}[\theta_t]
\end{aligned}
$$

which means $\widehat{\theta}_t$ is biased downward relative to population average $\mathbb{E}[\theta_t]$.

C.2 Cross-Sectional Estimation

Suppose $\theta_t = \kappa_t + b_t g(X)$ where $g(X)$ is a known function of observable X. We want to estimate κ and b. For simplicity, ignore unobserved heterogeneity in θ and drop constants from both first and second stage equations.

$$
\begin{aligned}
z' &\equiv \begin{pmatrix} S & Sg(X) \end{pmatrix} \\
x' &\equiv \begin{pmatrix} r_T & r_T g(X) \end{pmatrix} \\
y_t &\equiv r_t
\end{aligned}
$$

$$
\widehat{\theta}_t \underset{p}{\rightarrow} \mathbb{E}\left[zx'\right]^{-1} \mathbb{E}[zy_t]
$$

$$= \begin{pmatrix} \mathbb{E}\left[Sr_T\right] & \mathbb{E}\left[Sr_T g\left(X\right)\right] \\ \mathbb{E}\left[Sr_T g\left(X\right)\right] & \mathbb{E}\left[Sr_T g^2\left(X\right)\right] \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}\left[Sr_t\right] \\ \mathbb{E}\left[Sr_t g\left(X\right)\right] \end{pmatrix}$$

$$= \frac{\begin{pmatrix} \mathbb{E}\left[SVg^2\left(X\right)\right] & -\mathbb{E}\left[SVg\left(X\right)\right] \\ -\mathbb{E}\left[SVg\left(X\right)\right] & \mathbb{E}\left[SV\right] \end{pmatrix} \begin{pmatrix} \mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SV\right] \\ \mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SVg\left(X\right)\right] \end{pmatrix}}{\mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]^2}$$

$$= \frac{\begin{pmatrix} \mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]\mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SVg\left(X\right)\right] \\ \mathbb{E}\left[SV\right]\mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SVg\left(X\right)\right] - \mathbb{E}\left[SVg\left(X\right)\right]\mathbb{E}\left[\left(\kappa + bg\left(X\right)\right)SV\right] \end{pmatrix}}{\mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]^2}$$

$$= \frac{1}{\mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]^2} \begin{pmatrix} \kappa_t \left\{\mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]^2\right\} \\ b_t \left\{\mathbb{E}\left[SVg^2\left(X\right)\right]\mathbb{E}\left[SV\right] - \mathbb{E}\left[SVg\left(X\right)\right]^2\right\} \end{pmatrix}$$

$$= \begin{pmatrix} \kappa_t \\ b_t \end{pmatrix}$$

The above generalizes easily to vector valued b and g as in Section 5 where I use a vector of indicator functions for market equity quintile.

D Robustness

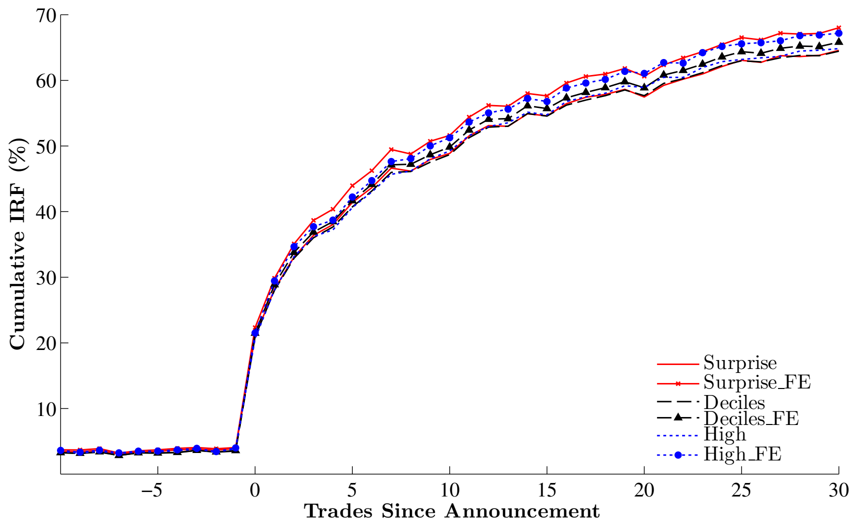
D.1 Choice of Instruments

Throughout this paper I use $z\left(S\right) = \left[\begin{array}{ccc} S & S^2 & S^3 \end{array}\right]'$ as the exogenous instruments and don't include time fixed effects in the second stage specification. This contrasts with the standard approach of sorting observations into discrete bins and forming portfolios. Differencing the "high" and the "low" portfolio returns removes time-variation in aggregate returns, similar to fixed effects. I reestimate the CIRF using alternative instruments, with and without fixed effects:

- As in DellaVigna & Pollet (2009), I sort earnings surprises each calendar quarter into deciles and construct a set of dummy variables for decile inclusion: $z\left(S\right) \equiv \text{Deciles} \equiv \left[\begin{array}{ccc} \mathbf{1}_{\text{Decile}=1} & \cdots & \mathbf{1}_{\text{Decile}=10} \end{array}\right]'$
- To avoid overfitting in the first stage, I construct the simple instrument $z\left(S\right) \equiv \text{High} \equiv \mathbf{1}_{\text{Decile}\geq 6}$ which is simply an indicator for the top half of the surprise distribution (each quarter)

Figure 17 shows the tick-by-tick impulse response function using the three sets of instruments. The estimated functions are quantitatively similar and yield the same qualitative conclusions.

Figure 17: CIRF Using Alternative Instruments



D.2 Overidentifying Restrictions and Linearity

Since I use overidentified IV, I examine the overidentifying restrictions for statistical and economic significance. Equation 1 states that $\mathbb{E}\left[CAR_t|\,S\right] = a_t + \theta_t \cdot \mathbb{E}\left[CAR_T|\,S\right] = a_t + \theta_t\left[\alpha + \gamma_1 S + \gamma_2 S^2 + \gamma_3 S^3\right]$. $\mathbb{E}\left[CAR_t|\,S\right]$ can also be directly estimated as $h_t\left(S\right) \equiv \mathbb{E}\left[CAR_t|\,S\right] = \alpha_t + \gamma_{1,t} S + \gamma_{2,t} S^2 + \gamma_{3,t} S^3$ (the cumulative reaction function at horizon t). This leads to the following overidentifying restriction[30]:

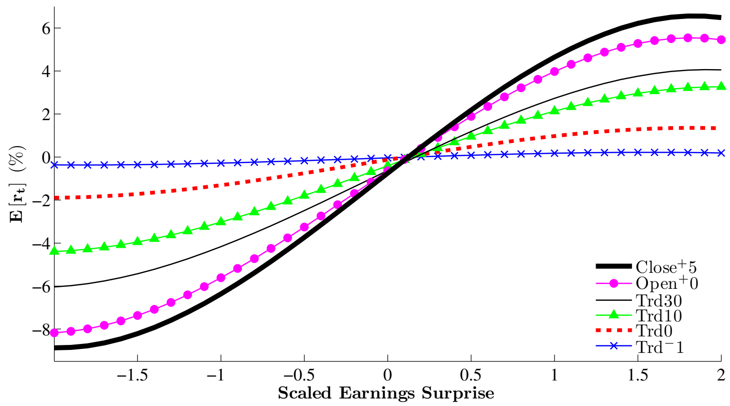
$\forall S, t:\ \alpha_t + \gamma_{1,t} S + \gamma_{2,t} S^2 + \gamma_{3,t} S^3 = a_t + \theta_t\left[\alpha + \gamma_{1,T} S + \gamma_{2,T} S^2 + \gamma_{3,T} S^3\right]$.

The restriction says that the cumulative reaction function at horizon t is proportional to the "long-horizon" cumulative reaction function at horizon T; the proportionality constant is equal to θ_t, the cumulative impulse response at horizon t [31]. A Hansen J-test does not reject the restriction at horizons shorter than Close$^+$0, by which point the cumulative impulse response is 90%. Figure 18 shows estimated functions $h_t\left(S\right)$ plotted from -2 to 2 standard deviations of S at various horizons.

Equation 1 also postulates a perfectly linear relationship between $\mathbb{E}\left[CAR_t|\,\Delta V\right]$ and ΔV. Some authors claim that arbitrageurs face short-sale constraints and this should slow the speed of price discovery for negative value shocks. This manifests in a positive cross-sectional mean value of $\mathbb{E}\left[CAR_t|\,\Delta V\right]$ since negative shocks are less fully incorporated than positive shocks and $\mathbb{E}\left[\Delta V\right] \approx 0$. At all horizons the second-stage intercept, a_t, is economically and

[30] Technically, the restriction is that $\mathbb{E}\left[\left(CAR_t - a + \theta_t \cdot CAR_T\right)\mathbb{S}\right] = 0$ where $\mathbb{S} \equiv \left[\begin{array}{ccc} S & S^2 & S^3 \end{array}\right]'$.
[31] Moreover, the reaction functions at *any* two horizons t_1, t_2 should be proportional with ratio $\theta^{0:t_1}/\theta^{0:t_2}$.

Figure 18: Cumulative Reaction Functions



statistically not different from zero, suggesting that $\mathbb{E}\left[CAR_t|\,\Delta V\right]$ is symmetric in ΔV (an odd function).

Furthermore, slower price adjustment to negative news means that $\mathbb{E}\left[CAR_t|\,\Delta V\right]$ is "steeper" for positive ΔV. I measure the impulse response as $\text{CIRF}_{\text{t,Neg}} = h_t\left(-2\right)/h_T\left(-2\right)$ and similarly $\text{CIRF}_{\text{t,Pos}} = h_t\left(2\right)/h_T\left(2\right)$. For $|S| > 2$ there is very little data and $h_t\left(\cdot\right)$ is poorly estimated. Figure 19a shows the estimated positive and negative impulse response functions as well as the pooled function (from Figure 7). The estimated functions are very similar, particularly for the first few trades post-announcement. As a second test, I estimate Equation 2 separately for positive and negative surprises. Given the reduction in sample size I estimate the first-stage using only the linear term, ignoring higher powers of S. Figure 19b shows the estimated functions; as in Figure 19a, the positive and negative impulse response functions are nearly identical for the first few trades. I do not find compelling evidence of non-linearity/asymmetry in the impulses response though these methods may lack power to detect such a phenomenon.

Figure 20: CIRF_t by Surprise Size

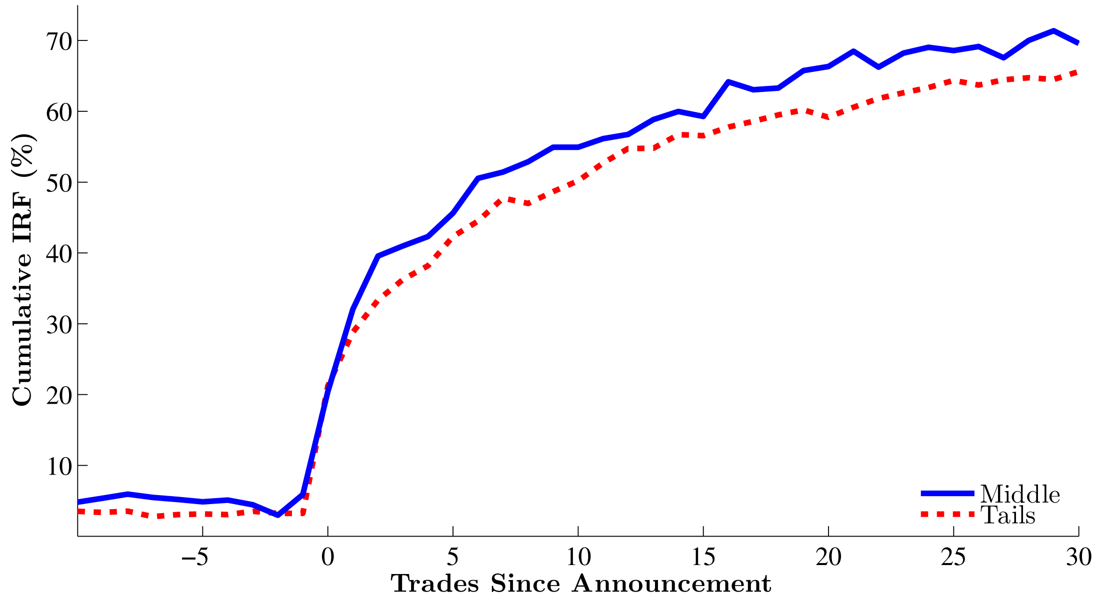
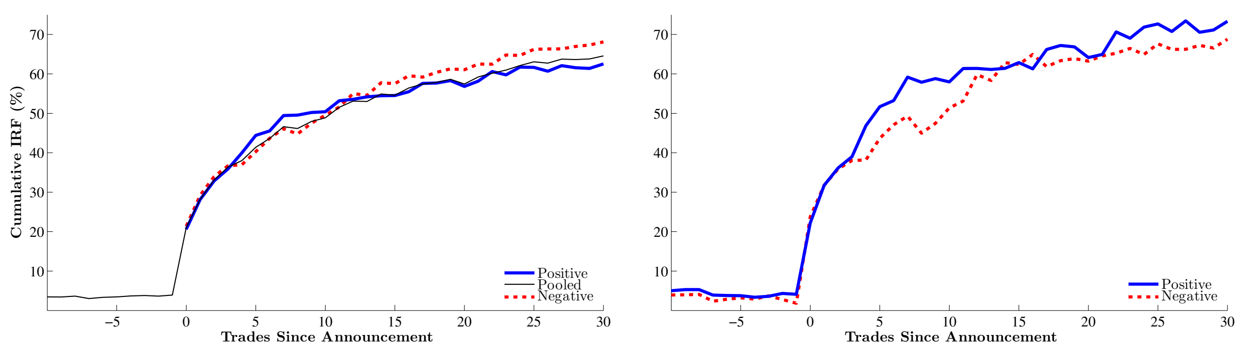


Figure 19: Symmetry of CIRF

(a) Using Reaction Functions

(b) Sorting on $Sign\,(S)$



Trade -1 is the last trade prior to the earnings press release. Trade 0 is the first trade post-announcement. Method1 estimates $\text{CIRF}_{t,Neg}$ as $h_t\,(-2)\,/h_T\,(-2)$ and similarly for $\text{CIRF}_{t,Pos}$. Method2 estimates Equation 2 separately for positive and negative earnings surprises except I use a linear first-stage.

As a further test of linearity I separately estimate CIRF_t for small and large announcements. Small announcements are those between the 25th and 75th percentiles of the surprise distribution, the "Middle". The remaining announcements are large, the "Tails". Figure 20 shows that CIRF_t is nearly identical for both small and large announcements. This means that speed of price discovery is independent of the size of the fundamental shock.

D.3 Firm Size

In Section 5 I measure CIRF_t as a function of firm size using the specification $\theta_t = \kappa_t + \mathbf{b}_t' \cdot \mathbf{1}_{\text{ME Quintile}}$. Appendix C shows that cross-sectional variation in θ_t can be specified as an arbitrary function of observable characteristics. In this section I use a continuous measure of size, $log\,(\text{MktCap})$, instead of quintile indicator functions. Table 13 repeats the analysis from Section 5 except that I specify $\theta_t = \kappa_t + b_t \cdot log\,(\text{MktCap})$. As in Table 7, $b_{\text{Open+0}}$ is significantly positive indicating that small firms have lower opening CIRF when estimated using all observations. The second panel of Table 13 shows that when measured in trade time, small firms incorporate news faster than small firms though not significantly so. The last panel shows that after conditioning on $\text{Trd}_{+30} \neq \varnothing$, the relationship between size and speed disappears. Figure 21 shows $\theta_{\text{Open+0}}$ as a function of firm size using the continuous and discrete specifications; the two methods yield similar estimates.

Table 13: Firm Size

Each column measures the cumulative impulse response (in %) by market equity quintile at the indicated horizon. $\theta_t = \kappa_t + b_t \cdot log\,(\text{MktCap})$. The first panel is estimated using all announcements. The second is trade by trade. The third restricts to announcements where $\text{Trd}_{+30} \neq \varnothing$. Standard errors are adjusted for heteroskedasticity. Absolute t statistics in parentheses.

	Open^{+}0	9:30	10:00	10:30	12:30	Close^{+}0
κ	17.7	32.4	48.4	51.2	61.0	60.7
	(1.5)	(2.7)	(4.0)	(4.2)	(5.2)	(5.5)
b_{MktCap}	8.6***	7.4***	6.2***	6.0***	5.0***	5.2***
	(4.9)	(4.1)	(3.4)	(3.3)	(2.9)	(3.1)

	Trd_{-1}	Trd_{+0}	Trd_{+10}	Trd_{+15}	Trd_{+20}	Trd_{+30}
κ	−3.5	25.1	79.6	85.4	82.2	97.5
	(0.6)	(2.4)	(4.5)	(4.4)	(4.2)	(4.6)
b_{MktCap}	1.0	−0.6	−4.2*	−4.2*	−3.3	−4.5*
	(1.4)	(0.4)	(1.9)	(1.7)	(1.3)	(1.7)

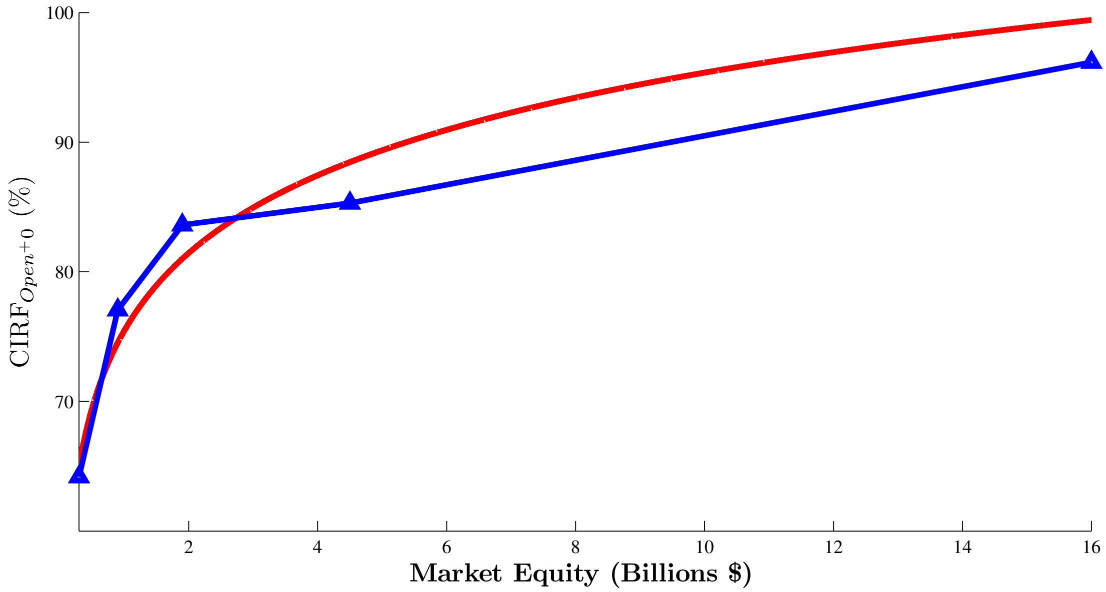
	Open^{+}0	9:30	10:00	10:30	12:30	Close^{+}0
κ	87.5	89.1	90.9	98.0	109.8	114.5
	(4.1)	(4.2)	(4.7)	(5.1)	(5.9)	(6.3)
b_{MktCap}	0.1	0.0	0.1	−0.7	−1.8	−2.2
	(0.0)	(0.0)	(0.0)	(0.3)	(0.8)	(0.9)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D.4 Value and Growth

Some previous studies find differences in post-earnings announcement drift for value and

Figure 21: Firm Size, Continuous and Discrete



growth stocks. Figure 22 shows no difference in CIRF_t.

Figure 22: CIRF_t by Book-to-Market