NORTHWESTERN UNIVERSITY

The Effect of Algorithmic Trading on Voluntary Disclosure

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Accounting Information and Management

By

Andrew Perry Stephan

EVANSTON, ILLINOIS

December 2017

ABSTRACT

The Effect of Algorithmic Trading on Voluntary Disclosure

Andrew Stephan

Algorithmic trading (AT) has grown dramatically in recent years and now makes up over half of all trades and orders in the market. I investigate whether and how AT affects voluntary disclosure by managers. I hypothesize that AT's differential ability to process information and its speed of trading affects how closely the market price aligns with the firm's fundamental value, and thereby changes the equilibrium disclosure level for managers. First, I investigate whether AT affects the quantity of voluntary disclosure from managers. Second, I examine why AT has such an effect. Last, I explore whether AT affects other disclosure characteristics such as specificity, readability, and timeliness. This question is important to academics, regulators, and the investing public due to the current debate over the desirability of algorithmic and high frequency trading in capital markets.

In Chapter 2, I define algorithmic trading, discuss its role in the capital markets, and review the prior literature examining its effects on price informativeness. AT utilizes many diverse strategies in the market, including market making, trading on statistical arbitrage opportunities, trading quickly in response to news, and executing large orders at efficient prices. The prior literature finds AT prices known information more efficiently, but it may reduce the amount of information known, leading to a net decrease in informativeness. In this chapter, I also review the literature regarding managers' forecast decisions.

In Chapter 3, I investigate the effect of AT on the quantity of voluntary disclosures, as proxied by management forecasts. I find the level of AT leading up to an earnings announcement is positively associated with the likelihood of issuing at least one forecast, the quantity of forecasts issued, and the number of days on which the firm issues forecasts. I utilize the introduction of the NYSE Autoquote, which exogenously increased AT for NYSE firms in 2003, as a quasi-natural experiment. I find results consistent with my main test, namely that NYSE firms see a greater increase in forecasts post-treatment compared to a matched sample of NASDAQ and AMEX control firms.

In Chapter 4, I examine why I find a positive association between AT and disclosure. Prior literature suggests that AT may decrease price informativeness by reducing the amount of information acquired by investors (Weller 2016). Under this explanation, managers would disclose more to offset this reduced information acquisition. I perform three tests to determine whether AT reduces the amount of information acquired by investors. First, I examine management forecast response coefficients, finding returns are more strongly associated with forecast surprises when AT leading up to the forecast issue date is higher. This is consistent with the market being less informed when the management forecast is released. Second, I test information acquisition directly, finding AT is negatively correlated with both EDGAR downloads and the magnitude of analyst forecast revisions. Moreover, AT is positively associated with the cost of informed trading, which may explain the decrease in information acquisition. Last, I examine whether the type of AT that improves price informativeness has the opposite effect on disclosure, and find that it does.

In Chapter 5, I investigate the association between AT and other disclosure characteristics. I find mixed results regarding AT and the specificity of disclosures. AT is negatively associated with forecast specificity, where point forecasts are the most specific, closed range forecasts are next, and open range forecasts are the least specific. For closed range forecasts, however, AT is negatively associated with the absolute range of the forecast, indicating these forecasts are more specific. I find no association between AT and the timeliness of disclosures and the readability of disclosures.

ACKNOWLEDGEMENTS

I am grateful for the invaluable support and guidance from my committee chair, Beverly Walther. I am also appreciative of the substantial feedback and advice from my committee members, Robert Magee, Robert Korajczyk, and Matthew Lyle.

I thank my professors at Kellogg who have provided me guidance and assistance throughout the doctoral program, including Daniel Aobdia, Craig Chapman, Ronald Dye, Danqi Hu, James Naughton, Tjomme Rusticus, Sri Sridharan, Linda Vincent, Clare Wang, and Wan Wongsunwai. I also thank my fellow doctoral students for their friendship and support, including Riddha Basu, Blake Bowler, Davide Cianciaruso, Pedro Gomez, Rita Gunn, Mark Kim, Brandon Lock, Rahul Menon, Spencer Pierce, Chis Rigsby, Kirti Sinha, and Ira Yeung. I am grateful to Lorraine Cassis, Mariela Flores, Susan Jackman, Tony Liu, Bradford Lund, and Joan Palmer for their help and assistance throughout the program.

The dissertation has benefited from valuable comments from Santhosh Ramalingegowda, Bobby Stoumbos, Laura Wellman, and workshop participants at Emory University, George Washington University, Indiana University, Southern Methodist University, the Trans-Atlantic Doctoral Conference, the University of Illinois at Chicago, the University of Colorado, and the University of Minnesota.

I gratefully acknowledge the generous financial support from the Kellogg School of Management at Northwestern University. I recognize Thomson Reuters for providing I/B/E/S analyst earnings forecast data.

Lastly, I would like to thank my family for the support and encouragement they have provided throughout the doctoral program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Algorithmic trading (AT)—the use of computer algorithms to execute trading strategies by automating the submission, cancellation, and execution of orders—has dramatically reshaped financial markets over the last 20 years, surpassing 50% of all traded volume by 2012 (Hendershott et al. 2011; Goldstein et al. 2014).[1] I investigate whether managers change their voluntary disclosure decisions in response to this new class of traders and investors. I find evidence consistent with AT reducing the amount of information acquired by non-algorithmic traders, leading managers to increase their voluntary forecasts because the market is less informed.

AT is a collection of diverse trading strategies linked by their ability to process vast quantities of data extremely quickly and trade at speeds faster than humans can process. These algorithms are necessarily restricted to machine readable data and pre-programmed algorithms. AT strategies differ in the type of information on which they trade. Some algorithms trade on new information that would not otherwise be priced. Statistical arbitrage strategies gather information from correlated price movements in other stocks, options, and futures that may be too costly for non-algorithmic traders recognize. Likewise, event arbitrage strategies, or algorithms that trade on released news, can reduce the effects of inattention around earnings announcements and more efficiently impound earnings into prices (Chakrabarty et al. 2016).[2]

---

[1] The estimate provided by Goldstein et al. (2014) is only for high frequency trading (HFT), a subset of AT.
[2] These trading strategies are not unique to AT, but the speed and information processing capability of AT have created a division between AT traders and non-AT traders. Thus, the implementation of the strategies by AT may influence disclosure strategies separately from those of non-AT traders.

If AT strategies such as these increase the amount of information known by the market, I expect that managers will disclose less. Under the expectations-adjustment hypothesis, managers forecast to align market expectations with their own private information (Ajinkya and Gift 1984; King et al. 1990). But given a cost to disclose, managers only do so when the market's expectations and the manager's private information are sufficiently misaligned (Verrecchia 1990). If AT increases the informativeness of prices, the market price will more closely reflect the private information of the manager, and voluntary forecasts will be less frequent.

Other AT strategies, however, trade on the order flow of informed traders, increasing the cost of informed trading and discouraging information acquisition. Theoretical models predict AT profits at the expense of non-algorithmic traders (Hoffman 2014; Yang and Zhu 2016). For example, order anticipation strategies predict incoming orders and trade ahead of them, making informed trading costlier (O'Hara 2015; Hirschey 2016; Korajczyk and Murphy 2016). If non-algorithmic traders see reduced gains from trade, they may choose not to become informed and instead leave the market (Weller 2016). In this case, market prices will less closely reflect the value of the firm as known by the manager, forecasts will be more informative, and I expect to observe a positive association between AT and disclosure.

Given that the prior literature is mixed on whether AT increases or decreases price informativeness, the collective effect of AT on disclosure is not clear, *ex ante*. Moreover, it is possible that AT will have no effect on disclosure for four reasons. First, the effects of AT on price informativeness may be too short-lived. Many AT strategies are executed over the course of minutes or hours, and if the effect of AT on price informativeness has a similar duration, it is less likely to affect disclosure decisions. Second, the amount of information known by the market may not be observable to the manager. I assume that the level of AT is too costly for most managers to acquire directly, but rather they compare the market price to their private information to estimate the informativeness of a given disclosure. If the manager cannot infer what the market knows from the price, then AT may have no effect on disclosure. Third, disclosure costs may vary with the amount of information known by the market. If a more informed market decreases the costs of disclosure, then an increase in price informativeness may be associated with increased disclosure. Fourth, the effects of different types of AT may be offsetting, leading to no change in price informativeness and no change in disclosure.

In Chapter 2, I define algorithmic trading, discuss its role in the capital markets, and review the prior literature examining its effects on price informativeness. The prior literature has found mixed results with respect to the effects of AT on price informativeness. AT appears to price known information more efficiently, but it may reduce the amount of information known, leading to a net decrease in informativeness. In this chapter, I also review the literature regarding managers' forecast decisions.

In Chapter 3, I examine AT's effect on the propensity to issue voluntary guidance and the quantity of such guidance issued. In my primary analyses from 2012 through 2016, I find the level of AT, measured in the quarter prior to an earnings announcement, is positively associated with both the likelihood of issuing guidance and the quantity of guidance issued on and following the earnings announcement. The results are robust to numerous specifications, including firm fixed effects and first differences models. Given the potential endogeneity of AT, I perform two tests to provide support for a causal interpretation of my results. First, I utilize the implementation of Autoquote on the NYSE in early 2003 as an exogenous increase to AT (Hendershott et al. 2011).

I find treated NYSE firms see a larger increase in voluntary guidance after Autoquote than untreated NASDAQ/AMEX firms, consistent with my primary analysis. Second, given the lack of an appropriate natural experiment during the time frame of my primary analyses, I utilize inverse propensity weighting and find consistent results.

As I discuss above, one possible explanation for the positive association between AT and disclosure is that AT reduces the amount of information known by the market by discouraging investors from becoming informed. Because I cannot measure the amount of information known by the market directly, in Chapter 4 I test four predictions that would be consistent with this explanation: 1) forecasts are more informative when AT is high, 2) AT reduces information acquisition, 3) informed trading is more costly when AT is high, and 4) the type of AT that improves price informativeness has the opposite effect on disclosure.

I first show that the market reaction to a given management forecast surprise is stronger when AT prior to the forecast is high. This suggests that the market knows less, as prices react more strongly to the forecast news. Second, I directly test the association between AT and information acquisition. I proxy for information acquisition using the number of SEC filing downloads from the EDGAR database and analyst forecast revisions. I find AT is negatively associated with non-robot EDGAR downloads and the magnitude of analyst forecast revisions, which is consistent with AT reducing information acquisition by traders. Third, I find AT is positively associated with the cost of informed trading, which may explain why investors acquire less information. Last, I identify two scenarios where the effect of AT on disclosure may be mitigated or in the opposite direction. The prior literature finds AT more efficiently prices earnings and news, which may lead to greater price informativeness. I show that AT surrounding the prior quarter's earnings announcement, of which I expect a greater proportion to be informativeness improving AT, is negatively correlated with the current quarter's guidance. I also find that the effect of AT on disclosure is mitigated when the volume of news during the quarter is high. These results are consistent with some AT improving the informativeness of market prices and leading to less disclosure.

Given the diversity of AT strategies, in Chapter 4 I also examine what type of AT is associated with increased disclosure. Prior literature finds active AT is more strongly associated with changes in price efficiency and information acquisition as compared to passive, market making AT (Brogaard et al. 2014; Weller 2016).[3] I observe, however, both active and passive AT are associated with increased disclosure. This surprising result suggests that market making AT, despite improving observed liquidity, may decrease the gains of informed traders.

In Chapter 5, I examine aspects of disclosure beyond the quantity of guidance issued. The effect of AT on the specificity of disclosure is mixed. Firms with high AT issue less specific guidance, but conditional on issuing a closed range forecast, the magnitude of the range is smaller for high AT firms.[4] I find no association between AT and the timing of firm disclosures or the readability of disclosures.

These questions are important because academics, regulators, and the public have questioned the benefits of AT for investors. In the media and popular press, AT and high frequency trading (HFT), a subset of AT, generated negative headlines for their contribution to the "flash

---

[3] Active trades are those that cross the mid-point of the bid-ask spread to execute an order (often these are market orders to trade a given volume at the best possible price). These are considered liquidity "taking" orders. Passive orders are offers to buy or sell at a given price that do not immediately cross the spread and execute, but rather sit on the order book until an active order trades against them. Passive orders are considered liquidity "making" orders.

[4] The most specific guidance is a point estimate, followed by a closed range (i.e., two values are given and the outcome is expected between them), and last an open range (i.e., the expectation is above or below a given value).

crash" in 2010 (Kirilenko et al. 2015) and the revelations of HFT front-running in Michael Lewis's *Flash Boys*.[5] Regulators have questioned the fairness of HFT, proposing new rules[6] to tighten the monitoring of such firms and prosecuting HFT broker-dealers for manipulative strategies (SEC Concept Release 2010).[7] The academic literature, however, finds it improves liquidity and price efficiency (Hendershott et al. 2011; Hasbrouck and Saar 2013; Brogaard et al. 2014; Conrad et al. 2015). I extend the prior literature that investigates the effects of AT on market quality and price informativeness. Weller (2016) finds that investors acquire less information when AT is high, leading to less informative prices. Because my results suggest that firms respond by increasing voluntary disclosures, the net effect of AT on the information environment may be mitigated or even positive.

I contribute to the academic literature in two additional ways. First, whereas prior research focuses on the relation between AT and market outcomes, to my knowledge I am the first to examine whether AT has real externalities with regards to management decision making. Beyond the mandatory disclosure regime set by regulators, firms have the option to disclose supplemental information in the form of forecasts, guidance, conference calls, press releases, and the text of mandatory disclosures such as 10-Ks. These firm disclosures reduce information asymmetry between the firm and investors, allowing a better allocation of capital and decreasing agency costs (Beyer et al. 2010).[8] It is important, therefore, to determine whether the rise in AT has positively

[5] Recent discussion of AT and HFT in the media has covered the IEX, a new stock exchange. The IEX deliberately slows orders by 350 microseconds to prevent HFT from trading ahead of incoming orders on other exchanges.
[6] See Regulation Automated Trading from the CFTC and a proposed rule by the SEC that would require many HFT broker dealers to register under FINRA.
[7] See http://www.sec.gov/News/PressRelease/Detail/PressRelease/1370543184457 for the SEC's press release regarding the prosecution of Athena Capital Research.
[8] I acknowledge there may be adverse effects to voluntary disclosure, such as increased short-termism by managers. As such, the documented increase in disclosure associated with AT is not necessarily good.

or negatively affected voluntary disclosures and the information asymmetry between firms and investors.

Second, my study contributes to the voluntary disclosure literature by examining how market microstructure affects managerial disclosure decisions. Prior literature has identified numerous factors that contribute to firms' voluntary disclosure decisions, including, but not limited to, upcoming equity and debt raises (Healy et al. 1999), insider trading (Cheng and Lo 2006), stock price changes (Sletten 2011), and the firm's investor base (Boone and White 2015). I contribute by identifying a new factor, the level of AT, that influences disclosure. Given the magnitude of AT in the market, it is important to understand its effect on voluntary disclosures and the information asymmetry between the firm and investors.

Further, my result is interesting because it differs from some predictions in the prior theoretical literature. Zhang (2001) suggests private information acquisition leads to greater information asymmetry between investors, which raises a firm's cost of capital. In response, managers are expected to disclose more to reduce their cost of capital. I, however, find managers respond to a decrease in private information acquisition with greater disclosure, suggesting managers are primarily concerned with the information asymmetry between the firm and investors, instead of the information asymmetry between investors.

CHAPTER 2

Literature Review

2.1 Algorithmic Trading Overview

AT uses computer algorithms to execute trading strategies by automating the submission, cancellation, and execution of orders (Hendershott et al. 2011). Over the last 20 years, AT has grown immensely, currently making up more than half of the trades and orders in our markets (Goldstein et al. 2014). Its growth, along with regulatory changes, has dramatically altered the landscape of equity markets.[9] The US now hosts 12 equity exchanges and more than 50 alternative trading systems (e.g., dark pools and crossing networks) (O'Hara 2015). With the advent of these new markets, trading activity has become increasingly decentralized; the share of volume of NYSE listed stocks traded on the NYSE dropped from 80% in 2003 to 25% in 2011 (Menkveld 2014). Algorithms are increasingly needed to navigate this more complex market environment.

AT generated headlines when a trading algorithm triggered the "flash crash" of 2010, executing a large sell order of the E-Mini S&P, resulting in a sharp market decline (Kirilenko et al. 2015).[10] The release of Michael Lewis' *Flash Boys* caused additional concerns that some traders possessed a technological advantage over others. This type of trading has grown so massive so quickly that regulators, academics, and practitioners are still trying to understand its effects on our capital markets.

---

[9] Regulation of Alternative Trading Systems (ATS) in 2000 allowed the growth of electronic communication networks to function as alternative trading venues. Regulation National Market System (NMS) in 2007 enacted new rules for the execution of orders across the various exchanges (O'Hara 2015). The latter is especially credited with the rise of HFT.

[10] According the CFTC-SEC Staff Report regarding the flash crash, a sell order from a mutual fund complex for the E-Mini S&P 500 futures contract was executed by an automated algorithm. The sell order consumed the available liquidity, causing prices to drop. Prices fell further as HFTs added additional selling pressures.

The finance literature has examined the effects of AT and HFT on the quality of financial markets, generally finding positive effects (Hendershott et al. 2011; Hasbrouck and Saar 2013; Brogaard et al. 2014; Conrad et al. 2015). Recent work has shown AT improves price efficiency in the short run; liquidity-taking HFT trades in the direction of permanent price changes and opposite transitory pricing errors (Brogaard et al. 2014). AT also effectively prices hard information releases extremely quickly, including macroeconomic news, shocks to futures and volatility indices, and earnings announcements (Zhang 2012; Scholtus et al. 2014). With the rise of HFT market makers, AT has narrowed spreads, increased depth, and reduced adverse selection (Hendershott et al. 2011; Hasbrouck and Saar 2013). These papers suggest markets are more efficient with greater AT; prices more accurately reflect aggregate expectations, and trading costs are reduced due to improved liquidity.[11]

There is concern, however, that the effects of AT are not uniformly positive. In addition to the flash crash, the academic literature finds HFT may impose greater adverse selection costs on slow traders (Chaboud 2014; Hoffman 2014), and in some cases may raise spreads and lower liquidity (Malinova et al. 2013; Menkveld and Zoican 2015). The negative effects of AT can even be intentional; some HFTs cause volatility and adverse selection through quote stuffing and spoofing (SEC Concept Release 2010; Egginton et al. 2014). While the average effects of AT on market quality may be positive, these cases reveal the potential negative consequences of AT.

[11] Prior studies generally test AT (which includes HFT) or HFT, but not both. The results across studies using AT and HFT appear consistent (e.g., the studies on the effects of AT/HFT on liquidity).

2.2 Algorithmic Trading and Price Informativeness

Although AT is made up of many diverse algorithms, they can be categorized into five broad groups: order anticipation, statistical arbitrage, event arbitrage, market making, and order execution. In this section, I review what these strategies do and how they might affect price informativeness.

Order anticipation strategies are designed to predict large incoming orders based on current order flow. Yang and Zhu (2016) model an informed trader and a "back-runner", who only observes the informed trader's period 1 order flow and then competes with the informed trader in period 2.[12] The effect is delayed price discovery; period 1 prices are less efficient because the informed trader randomizes their order flow to hide their private information. In period 2, however, prices are more efficient due to the addition of the back-runner trading. Hirschey (2016) finds evidence that HFTs lead non-HFT trades, especially when non-HFTs are less focused on disguising their order flow. He concludes that these strategies may increase the cost of informed trading for non-HFTs. If this is the case, AT may decrease price informativeness by reducing the likelihood a trader chooses to acquire costly information. It is possible, however, that under a price formation process where not all private information is impounded into prices, order anticipation strategies improve the informativeness of prices by preventing informed traders from hiding their private information.

Statistical arbitrage uses correlations between stock prices, orders, futures, indices, and currencies to predict price movements. If historically a price uptick in stock A is followed by a

---

[12] Many papers refer to the back-runners as front-runners. Front-running originally referred to market makers illegally executing orders on their own accounts before executing a customer's order. Back-runners, as defined in Yang and Zhu (2016), only trade on publicly observable order flow information.

stock uptick in price B, statistical arbitrage strategies will price the news for stock B more efficiently. Some of these complex correlations may be too costly to monitor and execute without algorithms. As such, I expect statistical arbitrage AT to increase price informativeness.

Event arbitrage AT trades in the direction of news, like non-AT traders, but processes the news automatically and trades quickly. Prior empirical literature shows AT prices hard information releases extremely quickly. Zhang (2012) utilizes trade data from NASDAQ that specifically identifies trades from HFTs and examines the stock market reaction to hard news (futures returns shocks and VIX returns shocks) and soft news (news articles). Zhang finds that HFTs dominate trading in response to hard news, reacting strongly within ten seconds and relinquishing their position within two minutes. Non-HFTs, however, are the primary traders in response to soft information. Scholtus et al. (2014) find similar results using message activity and fleeting orders as proxies for algorithmic activity and macroeconomic news announcements as events. They find a delay of even 300 milliseconds significantly reduces the profitability of trading on these news announcements. Rogers et al. (2016) document that some traders receive Form 4 insider trading filings up to 30 seconds early from an SEC subscription, and this substantially reduces the profitability of trading on the filing at the public release. Event arbitrage can have longer term effects beyond reacting to news within fractions of a second. Chakrabarty et al. (2016), using the same specifically identified HFT trades as Zhang (2012), finds HFT trades following an earnings announcement not only increases the speed at which the news is incorporated into prices, but also mitigate the delayed pricing effects of investor inattention. Overall, the evidence indicates event arbitrage improves price informativeness.

Some AT acts as market makers, dynamically supplying liquidity to the market. If AT increases (decreases) the noise in the supply of shares, prices will be less (more) efficient and reflect less (more) information known by traders (Grossman and Stiglitz 1980). Hendershott et al. (2011) use the introduction of the NYSE Autoquote as an exogenous increase in AT. They find that as firms utilize the Autoquote system, spreads narrow, adverse selection decreases, and trade-related price discovery also decreases. This suggests that at least over the short run, AT appears to improve the efficiency of prices. Conrad et al. (2015) find that prices more closely resemble a random walk when high frequency quoting is present. These results suggest market making AT improves price efficiency.

Despite this, recent evidence suggests that some AT market makers may employ order anticipation strategies to reduce the risk of adverse selection (Hasbrouck and Saar 2013; Korajczyk and Murphy 2016). Korajczyk and Murphy (2016) find that market making algorithms provide some liquidity at the beginning of a large parent order, but when they anticipate the additional shares to be traded, they turn and compete with the incoming orders. In this case, market making AT may increase the cost of informed trading, which would lead to a decrease in price informativeness. The net effect on price informativeness is not clear.

Last, traders use order execution algorithms to reduce the cost of executing large parent orders, which become expensive if they consume available liquidity or the market anticipates additional upcoming orders. This type of AT breaks up large orders into smaller pieces and executes the trades passively and actively at opportune times across various lit and dark markets to minimize the release of private information (O'Hara 2015). Effectively, this AT is designed to counteract other AT such as order anticipation strategies. Order execution AT reduces the cost of informed trading, which should encourage information acquisition and improve price informativeness, but it also hides the information held by an informed trader, which is expected to reduce priced information.

2.3 High Frequency Trading

HFT is a subset of AT and is commonly associated with the following characteristics: 1) extraordinarily high-speed programs for generating, routing and executing orders, 2) co-location, 3) short holding periods, 4) frequent submissions and cancellations of orders, and 5) ending the trading day in a flat position (SEC Concept Release 2010). Beyond these characteristics, HFT strategies generally differ from those of non-high frequency AT. Common HFT strategies include market making, order anticipation strategies, event arbitrage, and statistical arbitrage (Goldstein et al. 2014). Non-high frequency AT employs a different set of strategies, including, but not limited to, executing orders in dark pools, breaking up large orders into smaller pieces, earning the spread when trading, and even minimizing the effects of HFT when executing orders (O'Hara 2015).[13,14]

For purposes of this paper, I examine AT, including both high frequency and non-high frequency trading, for two reasons. First, the mechanism through which AT may affect voluntary disclosure, namely the amount of information incorporated into stock prices, could be driven by both HFT and non-high frequency AT. For example, if HFT order anticipation increases the amount of priced information, anti-HFT strategies by other algorithms may counteract this and hide the private information of informed traders.

---

[13] Degryse et al. (2015) find AT is negatively correlated with dark pool execution strategies, although some AT certainly uses dark execution.

[14] See O'Hara (2015) for an expanded discussion of non-high frequency AT trading strategies. ITG, a large broker-dealer, provides short descriptions of the AT strategies they employ at http://www.itg.com/marketing/ITG_Algo_ExecutionStrategies_Guide_20130701.pdf.

The second reason I study AT is that it is difficult to distinguish HFT and non-high frequency AT empirically.[15] Some papers make this distinction by specifically identifying trades by HFT broker-dealers, but these datasets are limited to either a small number of firms over a short time frame for a single exchange (e.g., the NASDAQ dataset used in Zhang (2012), Carrion (2013), and Brogaard et al. (2014)) or are only available for certain international exchanges (e.g., the ASX in Frino et al. 2015). The alternative to a dataset that identifies trader types is using market proxies for AT based on order submissions, executions, and cancellations (Hasbrouck and Saar 2013; Weller 2016). While these market proxies allow for longer sample periods and more firms, they do not distinguish between HFT and non-high frequency AT, as both employ high rates of order placement and executions, small trades, etc. I utilize the latter approach; given the effects of AT on voluntary disclosure are expected to be longer term in nature (measured in days instead of milliseconds), daily aggregate proxies for AT are sufficient to address my research question, and the expanded sample in both years and firms allows for improved causal testing through inverse propensity weighting.

2.4 Voluntary Disclosure

The expectations-adjustment hypothesis suggests that managers disclose to align market expectations with their own private information (Ajinkya and Gift 1984; King et al. 1990). There is a long literature examining the costs and benefits that determine these disclosure decisions. Disclosure costs can be direct (Verrecchia 1983) or indirect, such as proprietary costs (Dye 1986;

[15] The differences between HFT and non-high frequency AT can be uncertain and fluid. HFT of years past may be too slow to be considered high frequency today, as latencies have continued to decrease. Moreover, non-high frequency AT will often rapidly place and cancel large numbers of limit orders, similar to HFT. When the distinction between high frequency and non-high frequency can be measured in milliseconds or microseconds, high-frequency becomes a relative term. Both are substantially faster than human traders.

Wagenhofer 1990). The empirical evidence on the proprietary costs of disclosure is mixed; firms are more likely to redact information from mandatory disclosures when competition is high (Verrecchia and Weber 2006), but there is evidence of reduced voluntary guidance in low competition industries (Bamber and Cheon 1998).

The benefits of disclosure can accrue to numerous parties, including the firm, managers, and investors. The firm may benefit from disclosure by obtaining a lower cost of capital when raising funds (Korajczyk et al. 1991). Empirically we observe firms have higher levels of disclosure prior to issuing equity and debt (Lang and Lundholm 1993; Healy et al. 1999), and pre-IPO disclosures are also associated with less IPO underpricing (Schrand and Verrecchia 2005; Leone et al. 2007). Executives benefit by timing their insider trades around disclosures (or vice versa), selling shares following disclosures of good news and buying shares following disclosures of bad news (Noe 1999; Cheng and Lo 2006; Rogers 2008). Investors benefit from disclosure due to a reduction in information asymmetry; this results in improved liquidity (Healy et al. 1999; Leuz and Verrecchia 2000) and potentially better investment decisions (Beyer et al. 2010). The prior literature is mixed on whether disclosure benefits firms in the form of decreased litigation risk (Skinner 1997; Rogers and Van Buskirk 2009; Billings et al. 2015).

CHAPTER 3

Algorithmic Trading and the Quantity of Voluntary Disclosure

3.1 Hypothesis Development

My research question asks whether AT increases or decreases the quantity of voluntary disclosures made by management. The expectations-adjustment hypothesis suggests that managers voluntarily disclose to align market expectations with the manager's private information (Ajinkya and Gift 1984; King et al. 1990). If managers face some cost to disclose, one consideration in the disclosure decision will be how closely aligned the market price is with the manager's private information before disclosure. Consider a manager that obtains a private signal about the value of the firm and has the opportunity to disclose it at some cost.[16] The manager compares their signal to the current market value and determines whether the benefits of disclosing exceed the cost to do so. If AT improves the informativeness of market prices, prices will be less noisy and closer to the manager's private signal, on average. Thus, the manager will be less likely to disclose because the disclosure is less informative to an already informed market. If AT decreases the informativeness of prices, they will be noisier and further from the manager's private signal, on average. In this case, I expect the manager to be more likely to disclose, as the disclosure will be more informative and more likely to offset the cost doing so.

The theoretical model closest to the economic intuition I describe above is Verrecchia (1990). Verrecchia models a price maximizing manager that can disclose a private signal at some fixed cost. If the manager discloses, the market updates the price based on the precision of the

[16] The nature of this cost may vary, whether it is a proprietary cost of disclosure, the risk of missing a forecast, or an implicit commitment to continue forecasting in the future.

market's prior and the precision of the signal from the manager. The model predicts a negative association between the precision of the market's prior and the likelihood of disclosure. When the precision of the market's prior is high (low), a given disclosure will create less (more) of a change in the stock price, and therefore be less (more) likely to offset the cost of disclosure. The economic intuition in the model is the same; the more the market learns (i.e., the greater its precision), the less informative is a disclosure, and therefore managers are less likely to disclose. If AT affects how much information the market acquires and prices, I expect it to have an effect on disclosure.[17]

One important consideration is that managers may have different objective functions. Many models assume managers prefer to maximize firm value. In this case, as discussed in Verrecchia (1990), the association between price informativeness and disclosure is clear. Evidence suggests, however, that some managers may be motivated to disclose bad news, for instance to maintain a reputation of transparency (Teoh and Hwang 1991; Graham et al. 2005) or avoid potential litigation (Skinner 1994). The relation between price informativeness and the disclosure of bad news is not clear. On the one hand, if price informativeness is low, the stock price will not reflect the negative news known by the manager, and therefore managers may be more likely to withhold their bad news in the hopes of improvement in the future (Graham et al. 2005; Kothari et al. 2009). On the other hand, a market unaware about poor future performance may encourage managers to disclose to reduce litigation costs (Skinner 1994). As such, I do not

[17] Other disclosure models may predict an opposite association between how informed the market is and disclosure. In Dye (1998), information quality refers to the number of investors that learn whether the manager has received a private signal. When more investors are informed, managers with a bad signal cannot hide as well with firms that do not receive a signal, and therefore the threshold for disclosure is lowered and disclosure is more likely. In Penno (1997) high quality information is negatively correlated with the likelihood of receiving a signal, leading to less disclosure. Last, Zhang (2001) predicts greater information acquisition by traders increases adverse selection and the firm's cost of capital, leading managers to disclose more to reduce the information asymmetry between investors. In Chapter 4, I perform additional tests that suggest the Verrecchia (1990) model is the one most likely to explain my main result.

make a directional prediction for the relation between AT and disclosure for negative news, but rather test this relation empirically.

Three empirical papers provide evidence in support of the above intuition. Sletten (2011) finds an exogenous decrease in stock prices prompts managers to make good news disclosures. This finding supports the assumption that managers interested in maximizing their stock price only disclose when their news is sufficiently good, and the stock price is a reasonable metric to which managers compare their news. Li and Zhang (2015) observe an increase in short selling pressure increases price sensitivity to bad news and reduces the precision of bad news forecasts. Their paper provides evidence that managers consider price sensitivity to forecasts when deciding whether and what to disclose. Last, Balakrishnan et al. (2014) examine the exogenous loss of analyst coverage and find firms respond with increased voluntary disclosure. Analysts provide a public signal to the market that increases the precision of prices. When this signal is lost, firms substitute their own forecasts. Their finding is consistent with price informativeness affecting disclosure.[18]

Having established how price informativeness affects disclosure, I now consider how AT affects price informativeness. Based on the prior literature, I expect some AT to acquire and price new information (e.g., statistical arbitrage and event arbitrage, see Chapter 2.2). Other AT, however, may decrease the informativeness of prices by reducing the information acquired by non-

---

[18] My paper differs from the above by examining a new type of trading, AT, as opposed to information intermediaries such as analysts. AT comprises the majority of trades and orders in the market, and therefore it is important to understand its effect on disclosure. Unlike analysts, it is not obvious whether AT has improved or reduced price informativeness. As such, I cannot predict its effect on disclosure. Moreover, new information acquired by AT is private and revealed only through trading, which increases information asymmetry between investors, compared to analysts who publish their reports for a larger investing audience. Some models predict this will raise the firm's cost of capital and lead to increased disclosures (Zhang 2001). Therefore, it is not obvious, ex ante, whether managers are as likely to substitute their own disclosures for a loss of private information as they would for a loss of a public signal from analysts.

algorithmic traders. Finance theory predicts informed and slow traders are harmed when AT is high (Hoffman 2014; Yang and Zhu 2016). If traders profit less from their private information due to AT, then they are less likely to undergo costly information acquisition activities. Weller (2016) finds empirical evidence in support of this prediction, showing that a greater percentage of the cumulative abnormal returns in the month leading up to and including an earnings announcement are generated during the earnings announcement window when AT is high. His evidence suggests AT causes investors to reduce information acquisition activities by 8% to 29% in the month prior to an earnings release, leading to less informative prices.[19]

It is possible that AT will have no effect on disclosure for four reasons. First, the effects of AT on price informativeness may be too short-lived. Many AT strategies are executed over the course of minutes or hours, and if the effect of AT on price informativeness has a similar duration, it is less likely to affect disclosure decisions. Second, the amount of information known by the market may not be observable to the manager. I assume that the level of AT is too costly for most managers to acquire directly, but rather they compare the market price to their private information to estimate the informativeness of a given disclosure. If the manager cannot infer what the market knows from the price, then AT may have no effect on disclosure. Third, disclosure costs may vary with the amount of information known by the market. If a more informed market decreases the costs of disclosure, then an increase in price informativeness may be associated with increased disclosure. Fourth, the effects of different types of AT may be offsetting, leading to no change in price informativeness and no change in disclosure.

[19] I assume the effects of AT on prices persist for a sufficiently long time that the manager benefits from disclosing. Prior literature suggests this is reasonable; Weller (2016) finds differences in information acquisition in the month leading up to and including the earnings announcement, while Chakrabarty et al. (2016) find AT improves the efficiency with which earnings are priced over multiple days when attention is low.

The net effect of the different types of AT on price informativeness is not clear, and therefore I refrain from a directional prediction of the effect of AT on voluntary disclosure. As such, I state my hypothesis in the null form:

H1 (Null): The net effect of AT has no effect on the quantity of guidance issued.

3.2 Data and Research Design

3.2.1 Sample

My sample consists of all earnings announcements listed on I/B/E/S between February 1, 2012 and June 30, 2016.[20] I merge these announcements with CRSP and Compustat, requiring a stock price greater than $5 and a market capitalization greater than $10 million. I remove observations if the I/B/E/S earnings announcement date differs from the Compustat date by more than one calendar day. I require the prior earnings announcement date to be between 20 and 126 trading days prior to the current earnings announcement in order for the AT measurement window to be sufficiently long. I retain 34,310 firm-quarter observations (2,699 firms) over my sample period with available data.

3.2.2 Measuring AT

I obtain AT proxies from the SEC's Market Information Data Analytics System (MIDAS) dataset which collects data from 12 equity exchanges in the US through proprietary data feeds with microsecond timestamps.[21] The SEC provides summary data for volume, orders, trades, and

[20] The SEC MIDAS data is available beginning January 1, 2012. Eliminating earnings announcements prior to February 1, 2012 allows at least 31 days of measurement for the AT proxies.

[21] Following the flash crash in 2010, the SEC realized it did not have a system in place to effectively monitor and audit the US equity markets; the consolidated tape was insufficient due to its lack of odd-lot trades and missing order book data. As a result, the SEC developed MIDAS, requiring proprietary data feeds from the US equity exchanges. Data can be found at http://www.sec.gov/marketstructure/data.

cancellations by firm and day. This is the most complete source of market data for which AT proxies are available.[22]

I use four proxies for AT identified by the prior literature (Weller 2016). First, I calculate the odd-lot volume ratio (*OddLotVolumeRatio*) as the volume of trades executed in odd-lot sizes divided by the total volume traded, where greater odd-lot trades are associated with more AT (O'Hara et al. 2014). Second, I measure the trade-to-order ratio as the total volume traded divided by the total volume of orders placed. The trade-to-order ratio (*TradeToOrderRatio*) is negatively correlated with AT because algorithms place and cancel high numbers of orders when executing trades (Hendershott et al. 2011). Third, the cancellations-to-trades ratio (*CancelToTradeRatio*) is the number of orders cancelled divided by the number of trades executed and is positively correlated with AT, again because algorithms place and cancel large numbers of orders (Hasbrouck and Saar 2013; Hendershott and Riordan 2013). My final proxy is trade size (*TradeSize*), calculated as the total volume traded divided by the number of trades. Trade size is expected to be negatively correlated with AT, as algorithms execute a greater number of small orders to trade a given volume (Brogaard et al. 2014; Menkveld 2014; O'Hara 2015).[23]

I take the average of each proxy, measured daily by firm, beginning five days following the prior earnings announcement to two days before the current announcement ("preannouncement window"). I calculate the log of each proxy to determine the final variable. The means for each measure, presented in Table 1, are consistent with the prior literature (Weller 2016). Most

---

[22] MIDAS does not include trade data from the numerous alternative trading systems, but only from 12 US equity exchanges. Alternative trading systems make up between 10-15% of total trading volume on a given day (see https://www.sec.gov/marketstructure/research/alternative-trading-systems-march-2014.pdf).

[23] The NYSE and AMEX do not report cancellations to trades and odd lot volume. These measures are appropriately adjusted such that bias is not expected (i.e., the total number of trades in the cancellations to trades ratio only contains trades from the exchanges that report cancellations).

importantly, the proxies are correlated as expected (see Table 2); *OddLotVolumeRatio* and *CancelToTradeRatio* are positively correlated with each other (0.57), and negatively correlated with *TradeToOrderRatio* (-0.59 and -0.78, respectively) and *TradeSize* (-0.76 and -0.30, respectively). To obtain my final measure, *AT*, I perform a principal components analysis on *OddLotVolumeRatio*, *TradeToOrderRatio*, *CancelToTradeRatio*, and *TradeSize*. The first principal component has an eigenvalue of 2.58 and is positively correlated with *OddLotVolumeRatio* and *CancelToTradeRatio* and negatively correlated with *TradeToOrderRatio* and *TradeSize* as expected (see Table 2). Thus, I consider it an appropriate proxy for AT.

3.2.3 Research Design

The base model used to test the effect of AT on voluntary disclosure is as follows:

$$\begin{aligned} Guidance_{i,t} = \alpha &+ \beta_1 AT_{i,t} + \beta_2 UE_{i,t} + \beta_3 PosUE_{i,t} + \beta_4 NegUE_{i,t} + \beta_5 Loss_{i,t} \\ &+ \beta_6 CAR\_Pre_{i,t} + \beta_7 PPUE_{i,t} + \beta_8 LnMktCap_{i,t} \\ &+ \beta_9 AnalystFollowing_{i,t} + \beta_{10} Dispersion_{i,t} + \beta_{11} SalesGrowth_{i,t} \\ &+ \beta_{12} Turnover_{i,t} + \beta_{13} Volatility_{i,t} + \beta_{14} Skewness_{i,t} + \beta_{15} Spread_{i,t} \\ &+ \beta_{16} InsiderSales_{i,t} + \beta_{17} NewsVolume_{i,t} + \beta TimeFE + \varepsilon_{i,t} \end{aligned} \tag{1}$$

$Guidance_{i,t}$ represents the voluntary disclosure for firm i in quarter t. Guidance is obtained from the I/B/E/S Guidance dataset, which records firm forecasts for EPS, sales, EBITDA, EBITDA per share, capital expenditure, dividends per share, funds from operations, fully reported EPS, gross margin, net income, operating profit, pretax income, ROA, and ROE. I measure guidance in three ways. I create an indicator variable, *Guider*, equal to 1 if the firm issues any guidance in the [-1,

+1] trading day window around the current earnings announcement ("announcement window"), and 0 otherwise. I also use two continuous variables for the quantity of guidance, *GuideCount* and *GuideDays* (Billings et al. 2015). The former is a count of the pieces of guidance issued in the [-1, +1] day announcement window. For *GuideCount*, I consider each forecast to contain separate information. *GuideDays* is the number of distinct days on which the firm issues guidance, measured from the day prior to the current earnings announcement to two days prior to the next earnings announcement. In this case, I assume each day that guidance is issued reveals new information and multiple forecasts on the same day report the same information.[24]

As shown in Table 1, firms issue guidance during the disclosure period in 71% of firm-quarters. The mean (median) quantity of guidance issued is 2.22 (2), while the mean number of days on which it is issued is 1.04 (1). I find 598 firms (3,802 observations) never issue guidance, 1,096 firms (16,163 observations) always issue at least one forecast, and the remaining 1,005 firms (14,345 observations) issue at least one forecast in one quarter and zero forecasts in at least one quarter.

$AT_{i,t}$ is the proxy for AT as discussed in Chapter 3.2.2. The remaining covariates follow the voluntary disclosure models in Rogers and Van Buskirk (2013) and Billings et al. (2015). I control for earnings news with $UE_{i,t}$, $PosUE_{i,t}$, $NegUE_{i,t}$, and $Loss_{i,t}$. $UE_{i,t}$ is the standardized unexpected earnings calculated as the actual earnings per share less the mean consensus forecast, based on the most recent forecast from each analyst up to two trading days prior to the earnings announcement, scaled by the share price at the beginning of the preannouncement period (Billings

[24] AT is measured strictly prior to the measurement of guidance. For a given earnings announcement, I calculate AT in the quarter leading up to, but prior to the announcement, and I measure guidance during the announcement window. When calculating *GuideDays*, I require more than a three-day announcement window to provide additional opportunities to issue guidance, so I include the quarter following the announcement.

et al. 2015). If there is no analyst forecast for the quarter, I use the seasonal random walk earnings surprise. $PosUE_{i,t}$ and $NegUE_{i,t}$ are indicators equal to 1 if $UE_{i,t}$ is greater than +0.0001 or less than -0.0001, respectively, and 0 otherwise (Rogers and Van Buskirk 2013; Billings et al. 2015). $Loss_{i,t}$ is an indicator equal to 1 if the firm recorded negative earnings, 0 otherwise.

To account for news and performance contemporaneous with my AT measurement, I control for the cumulative abnormal return in the preannouncement period ($CAR\_Pre_{i,t}$) as estimated from the Fama-French three factor model. I also include $PPUE_{i,t}$, which is the proportion of the last four quarters for which the firm beat earnings expectations (i.e., the proportion for which $PosUE = 1$). I control for the firm's information environment with the log market value of equity ($LnMktCap_{i,t}$), the number of analysts following the firm ($AnalystFollowing_{i,t}$), and the standard deviation of the forecasts that make up the consensus earnings expectation ($Dispersion_{i,t}$). Contemporaneous news events during the quarter may both drive AT trading and cause firms to respond with greater disclosures. I control for news ($NewsVolume_{i,t}$) using the RavenPack aggregate event volume. This variable captures the volume of events related to a firm by measuring the quantity and novelty of news articles in the 91 days prior. For example, one event that generates five news articles will have a lower aggregate event volume than five events that generate one news article each. I measure *NewVolume* on the date closest but prior to the upcoming earnings announcement to capture the amount of news during the quarter in which I measure AT.

I also include control variables used in the Kim and Skinner (2012) litigation risk model. $SalesGrowth_{i,t}$ is the percentage change in sales compared to the prior quarter. $Turnover_{i,t}$, $Volatility_{i,t}$, and $Skewness_{i,t}$, measure the share turnover, return volatility, and return skewness, respectively, over the [-126,+5] trading day period around the prior quarter's earnings announcement ("market control window"). AT likely causes changes in turnover, volatility, and the skewness of returns, and thus I calculate these controls prior to the measurement of AT (Hasbrouck and Saar 2013; Malinova et al. 2013; Scholtus et al. 2014). I measure bid-ask spreads ($Spread_{i,t}$) in the market control window as well (Hendershott et al. 2011; Hasbrouck and Saar 2013).

I include insider sales, as these are associated with firm disclosures (Huddart et al. 2007; Cohen et al. 2012; Billings and Cedergren 2015). $InsiderSales_{i,t}$ is calculated as the net purchases and sales volume for officers and directors in the preannouncement period, scaled by the market value of the firm. Last, time fixed effects (quarterly) are included to account for changes in disclosure levels over time that may be correlated with AT.

Other variables that have been associated with disclosure, including industry (Francis et al. 1994), market competition, firm age (Chen et al. 2002), and investor sentiment (Bergman and Roychowdhury 2008) are controlled through time or firm fixed effects (see below for the discussion of firm fixed effects). I assume equity based incentives (Nagar et al. 2003), manager specific effects (Bamber et al. 2010), and non-manager employee ownership (Bova et al. 2015) will be sufficiently static over the 4.5 year sample period that they will not alter the results.

I present summary statistics for the above variables in Table 1. 58% of firms have positive earnings surprises, while 23% report a loss. The median analyst coverage for my sample is 10 analysts. Median sales growth is 5% and median returns in the quarter prior to an announcement are near zero. These descriptive statistics are consistent with the prior literature except for the percentage of guiding firms. I find a moderately higher percentage of firms are guiders than prior research, likely due to my recent sample period (2012-2016).

I consider three specifications of the above base model. First, I use a logit (*Guider*) or OLS (*GuideCount* and *GuideDays*) model as written. Second, I add firm fixed effects, as disclosure levels and *AT* can be persistent within a firm. Last, I run a first differences model where the guidance, *AT*, and controls are measured as the change from the prior quarter.

3.3 Results

I present the results of *H1* in Table 3. Columns (1), (2), and (3) in Panel A test *Guider*, *GuideCount*, and *GuideDays*, respectively, using the base model described in Chapter 3.2.3. I use a logit model when testing *Guider*, and an OLS model when testing *GuideCount* and *GuideDays*. I find *AT* is positively correlated with each measure of guidance, though is significant only for *Guider* and *GuideCount* (significant at the 1% and 5% levels, respectively). When *AT* is high prior to an earnings announcement, firms are more likely to issue guidance and issue more guidance.

To address the possibility that an omitted firm characteristic correlated with *AT* and disclosure induces these results, I add firm fixed effects to the prior model in columns (4), (5), and (6), finding stronger results. All three guidance variables are positively correlated with *AT* (significant at 1%). In terms of the economic magnitude of these effects, a one standard deviation increase in *AT* is associated with a 28.5% increase in the likelihood of issuing guidance. *GuideCount* increases by 0.07 (an 3.5% increase over the median) and *GuideDays* by 0.03 (a 3% increase over the median).

Table 3 Panel B presents the results of a first differences model, where each variable is calculated as the difference from the prior quarter. The change in *AT* is positively associated with a change in guidance (significant at 1%), whether measuring the likelihood of issuing guidance, the quantity issued, or number of days on which it is issued. Table 3 results are robust to double clustering the standard errors by firm and calendar quarter. As a whole, Table 3 suggests that AT is positively associated with the quantity of management forecasts.

3.4 Causal Inference

In this section, I identify three risks regarding a causal interpretation of my results and discuss how I mitigate them. The first risk is reverse-causality; that is, disclosure is causing AT. To address this concern, I measure AT strictly prior to the disclosure outcome variable and control for prior guidance. The results are unaffected (see Table 3 and Chapter 3.4.3). The second risk is an omitted firm-level correlated variable. To mitigate this possibility, I include firm fixed effects and estimate a first differences specification, both of which produce consistent results (Table 3).

The third risk is a time varying omitted correlated variable during the AT measurement window; e.g., a firm event or news story occurs prior to the earnings announcement that both causes AT to increase and managers to respond with greater guidance. To address this concern, I first control for other market variables during the AT measurement window that would reflect time-varying firm-level events, including share turnover, return volatility, market returns, and news volume (see Chapter 3.4.3). If there is an omitted variable during the measurement quarter causing the observed result, it would have to influence AT separately from the other market measures included. In the following subsections, I use two additional techniques to address this same issue. First, I identify a natural experiment, the introduction of Autoquote on the NYSE in 2003, that exogenously increases AT for a set of firms. Second, I utilize inverse-propensity reweighting to generate a sample weighted on the difference between the predicted level of AT and the observed level of AT.

3.4.1 NYSE Autoquote

In early 2003, the NYSE introduced the Autoquote system which automatically disseminated inside quotes whenever there was a relevant change in the limit order book. Prior to its introduction, the inside quote had to be manually disseminated by the NYSE specialists. As documented in Hendershott et al. (2011), this increase in the rate of quote dissemination was beneficial for algorithmic traders, but changed little for human traders. They document Autoquote is positively correlated with AT, and, using its introduction as a natural experiment, show AT causes liquidity to improve. I adopt the same setting in order to test the effect of an exogenous increase in AT on the quantity of guidance issued by firms.

Autoquote was phased in beginning January 29, 2003 and ending May 27, 2003. For my treatment group, I begin with a sample of all NYSE firms that have four earnings announcements in the 18 months prior to the first phase-in date and four earnings announcements in the 18 months following the last phase-in date. I keep the four earnings announcements pre- and post-treatment that are closest to the beginning and end of the phase-in, for a final sample of eight quarters per firm (four pre-treatment and four-post-treatment).[25] For a control group, I begin with all NASDAQ and AMEX firms, keeping the same eight quarters discussed above.

I propensity match my treated and control firms to ensure they are as similar as possible. For each firm, I average the variables used in the propensity model over the four pre-treatment quarters. I match by firm, instead of by quarter, to ensure each firm has four pre- and post-

[25] The phase-in window is excluded due to uncertainty of when firms receive the treatment.

treatment quarters in the final sample, and a full year of fiscal quarters is represented in each period. I model the propensity to be treated (i.e., be a firm listed on the NYSE) as follows:

$$\begin{aligned} Treated_i = \alpha &+ \beta_1 LnMktCap_i + \beta_2 TobinQ_i + \beta_3 UE_i + \beta_4 LnCAR_i + \beta_5 Volatility_i \\ &+ \beta_6 LnAnalystFollowing_i + \beta_7 GuideCountEPS_i + \beta_8 GuiderEPS_i \\ &+ \beta_9 GuideCount_i + \beta_{10} Guider_i + \beta_{11} ROA_i + \beta_{12} Spread_i \\ &+ \beta_{13} Turnover_i + \beta_{14} Price_i + \beta IndustryFE + \varepsilon_i \end{aligned} \tag{2}$$

Treated is equal to 1 if the firm is listed on the NYSE, 0 otherwise. I include in the model firm characteristics (*LnMktCap*, *TobinQ*, *LnAnalystFollowing*, *Price*), firm performance (*UE*, *ROA*), market controls (*Spread*, *Turnover*, *Volatility*), prior forecasting behavior (*GuideCountEPS*, *GuiderEPS*, *GuideCount*, *Guider*), and industry fixed effects. *LnMktCap* is the natural log of the market value of equity at the beginning of the quarter. *TobinQ* is the market value of the firm divided by total assets. *LnAnalystFollowing* is the natural log of one plus *AnalystFollowing*, as defined previously. *Price* is the average stock price for the firm over the pre-treatment period. *UE* is unexpected earnings, as defined previously. *ROA* is return on assets, measured as earnings divided by total assets. *Spread, Turnover,* and *Volatility* are the average bid-ask spread, share turnover, and volatility, respectively, over the pre-treatment period. *GuideCountEPS* is the count of the number of pieces of EPS guidance issued by the firm in the pre-treatment period. *GuiderEPS* is an indicator equal to one if the firm issues at least one EPS forecast in the pre-treatment period. *GuideCount* and *Guider* are as defined previously, measured over the pre-treatment period.

In both the propensity model and my later regressions, I examine guidance variables for EPS forecasts only and for all types of forecasts combined. I do so because at the beginning of 2003, near the Autoquote implementation date, I/B/E/S began collecting additional types of forecasts beyond EPS and sales. If this change in collection was not immediate or uniform, it may influence the results. By testing EPS only, I ensure the same guidance is being measured pre- and post-treatment.

Table 4, Panel A presents the results of the propensity model. I match NYSE firms to non-NYSE firms one-to-one without replacement, keeping the firm with the closest propensity score and requiring the score to be within a caliper of 0.05. This procedure results in 331 treated NYSE firms matched to 331 untreated NASDAQ or AMEX firms. In Table 4, Panel B, I test the covariate balance between the treated and untreated firms. All but three covariates are not significantly different between the treated and untreated groups in the pre-treatment period, most notably the propensity score and the four guidance variables. Three variables, *LnMktCap*, *TobinQ*, and *LnAnalystFollowing* are statistically different between the two groups. As such, I include these variables as controls in the final test.

I present the results of the differences-in-differences regression in Table 4, Panel C. *Treated* indicates the firm receives the Autoquote treatment as it is listed on the NYSE. *Post* indicates the earnings announcement falls after the end of the Autoquote implementation. The variable of interest is the interaction of *Treated* and *Post*. In both regressions, *GuideCount* in column (1) and *GuideCountEPS* in column (2), the coefficient on the interaction is positive and significant (at 10% and 5%, respectively). This implies that the quantity of disclosure increases more for treated firms that utilize Autoquote than untreated firms that do not.

One concern with a differences-in-differences test is that different pre-treatment trends can drive the results, even if the covariates are fairly well balanced. I address this issue by graphing the pre-treatment trends in *GuideCountEPS* and *GuideCount.* In Figure 1, I present the average *GuideCountEPS* per quarter from June 2000 through June 2006. The vertical black lines represent the beginning and end of the treatment period. We observe that although the pre-treatment trends are not perfectly parallel, they are not trending in opposite directions in a way that would explain the results. Especially in the periods immediately before treatment, the trends are similar. Figure 1 also indicates that post-treatment, the treated NYSE firms see an increase in forecasting greater than the untreated firms, and this increase persists over time. In Figure 2, I present the same chart using *GuideCount*. The results are consistent.

Overall, these results support my primary findings and help to address the concern that an omitted correlated variable drives the main result.

3.4.2 Inverse Propensity Reweighting

Because the Autoquote experiment tested above falls before the time period of my main tests (2012 through 2016), I implement inverse propensity reweighting with regression adjustment (IPWRA) to reduce the risk of an omitted correlated variable during my primary timeframe (Woolridge 2010). IPWRA weights observations according to the inverse of their probability to be a treated or control unit in order to estimate potential outcome means. For example, a treated unit with a high probability of being treated receives a weight close to 1, whereas a treated unit with a low probability of treatment receives a weight higher than 1 (Imbens 2000, Hirano et al. 2003). IPWRA addresses the concern of endogeneity by creating two samples with similar propensities for AT, but different observed levels of AT. This ensures there is no selection on observable characteristics, and reduces the risk of an omitted correlated variable. Intuitively, the method is similar to matching with replacement and is competitive with the most effective matching estimators (Busso et al. 2014). Because I have a continuous treatment variable, I divide AT into quintiles and run a multinomial logit propensity model.

I include in the propensity model covariates that are available at the time I measure AT:

$$\begin{aligned} AT_{i,t}^{Quintile} = {} & \alpha + \beta_1 LogMktCap_{i,t} + \beta_2 Volatility_{i,t} + \beta_3 Turnover_{i,t} + \beta_4 Spread_{i,t} \\ & + \beta_5 Skewness_{i,t} + \beta_6 AnalystFollowing_{i,t} + \beta_7 Dispersion_{i,t} \\ & + \beta_8 CAR\_Pre_{i,t} + \beta_9 PPUE_{i,t} + \beta_{10} InsiderSales_{i,t} + \beta TimeFE + \varepsilon_{i,t} \end{aligned} \tag{3}$$

The first five variables are expected to be correlated with AT (Groß-Klußmann and Hausch 2011; Hendershott et al. 2011; Hasbrouck and Saar 2013; Conrad et al. 2015). I also control for the information environment (*AnalystFollowing*, *Dispersion*), contemporaneous performance (*CAR_Pre*, *PPUE*), and insider sales (*InsiderSales*). See Chapter 3.2.3 for variable definitions.

IPWRA performs best when the overlap between treated and untreated samples is high (Busso et al. 2014). Best practice involves trimming the high and low estimated propensities to ensure the treated and untreated samples have sufficient overlap (Black et al. 2013). I estimate the propensity to be treated and trim the sample of observations with propensities to be treated less than 5% or greater than 95%. I then re-estimate the propensity to be treated and trim again. I repeat the estimation and trimming a third time to ensure proper overlap. In Figure 3, I graph the kernel density function of the propensity scores for a given firm quarter to be in AT Quintile 1 (i.e., low AT) before trimming. As expected, there are many AT Quintile 5 observations with very low propensities to be in AT Quintile 1. Meanwhile, AT Quintile 1 observations have very few observations with low propensities to be in AT Quintile 1. Given the low overlap, these low propensity observations are not optimal to compare. In Figure 4, I present the same kernel density plot after trimming. Given the propensities are re-estimated on the trimmed sample, there is still a spike in the AT Quintile 5 group in the low probability section, but it is substantially less than before trimming. The overlap appears better.[26]

The outcome model includes information released on or after the earnings announcement, along with firm size, performance, and analyst characteristics:

$$
\begin{aligned}
GuideCount_{i,t} & \\
&= \alpha + \beta_1 LogMktCap_{i,t} + \beta_2 CAR\_Pre_{i,t} + \beta_3 UE_{i,t} + \beta_4 PosUE_{i,t} \\
&+ \beta_5 NegUE_{i,t} + \beta_6 Loss_{i,t} + \beta_7 SalesGrowth_{i,t} \\
&+ \beta_8 AnalystFollowing_{i,t} + \beta_9 Dispersion_{i,t} + \beta TimeFE + \varepsilon_{i,t}
\end{aligned} \tag{4}
$$

The results, presented in Table 5, show that for firms with similar predicted levels of AT, firms with a higher realization of AT issue more voluntary guidance. Note that the comparison must be made for each quintile. For example, I first estimate the propensity to be in AT Quintile 3, then execute the IPWRA model comparing AT Quintile 3 firm quarters to other quintiles with similar propensities to be in AT Quintile 3. The 1 vs. 3 comparison in the AT Quintile 3 test indicates that for firms with a similar predicted level of AT, the firms with actual AT in the third quintile (greater AT) have higher average disclosure (1.34 more pieces of guidance issued,

[26] Untabulated tests show the results are robust (even stronger) if I do not trim the sample.

significant at 1%) than firms in the first quintile (less AT). Across the 20 comparisons, 19 are significant at the 10% level or better. The effects also appear stronger in the lower quintile comparisons, which may suggest the effect of AT on disclosure may not be linear. Overall, the IPWRA results appear consistent with my primary findings in Table 3.

3.4.3 Robustness Tests

I run multiple robustness tests (untabulated) on fixed effects model tested in Chapter 3.3. First, I include as controls the guidance count during the preannouncement window and the market control window to ensure that the firm's prior disclosure behavior is not correlated with both AT and the firm's future disclosures. The results are consistent with those presented in Table 3. Next, I include institutional ownership as calculated from Thomson, as institutions may use more AT and also affect firm disclosure decisions, and find similar results. [27]

One remaining concern is that unobservable information events during the preannouncement period both increase AT and cause firms to increase their disclosures in response. In my general model, I include *CAR_Pre* and *NewsVolume* to account for firm specific information events in the preannouncement window. Information, however, may manifest itself as additional turnover or volatility.[28] As a robustness test, I add volatility, turnover, skewness, and spread, measured contemporaneously with *AT* in the preannouncement period, to the model. I find results are robust to these additional controls. I also control for the absolute value of abnormal returns in the preannouncement period, instead of the current signed returns, finding consistent results.

[27] I exclude institutional ownership from my primary model because it reduces the sample size substantially.
[28] These variables are measured prior to AT in the base model because AT is expected to influence them directly, which may bias coefficients and affect inferences.

As I discuss in Chapter 3.1, it is not obvious the main result should be consistent across good and bad news disclosures. I test my primary model on subsamples of positive and negative earnings surprises. In both subsamples, the coefficients on AT are positive and are not significantly different from each other. This finding suggests the AT mechanism is consistent across good and bad news. I run a similar test conditional on whether the forecast issued is below expectations or not. To do so, I first identify the longest-horizon EPS forecast issued for each quarter and categorize each firm-quarter into two groups: forecasts in-line with or above expectations, and forecasts below expectations. I cannot run the analysis on the indicator, *Guider*, as the test requires the firm issue at least one EPS forecast. I test *GuideCount* and *GuideDays* on the good/bad news forecast subsamples. For both measures, I find the coefficient on AT is significant only for the good news forecasts and insignificant for the bad news forecasts. Overall, the results indicate the effect of AT may be stronger for good news disclosures due to managerial incentives, but at least for some managers, the effect of AT remains for negative news as well.

Managers may potentially change their disclosures in anticipation of future AT, as opposed to responding to prior AT. I replace my main AT variable with the next quarter's AT value, finding future AT is not significantly correlated with disclosure. Last, I perform my main analysis excluding the always-guiding firms and the never-guiding firms, and find no changes in the results.

3.5 Conclusion

In this chapter, I document a positive association between AT and the quantity of voluntary forecasts, including the likelihood of issuing a forecast, the number of forecasts issued, and the number of days on which a firm issues forecasts. The result is robust to numerous alternative specifications, including firm fixed effects, first differences, and inverse propensity reweighting.

Moreover, I find consistent results using the NYSE Autoquote implementation as a quasi-natural experiment that affects the level of AT.

CHAPTER 4

Why Does Algorithmic Trading Affect Disclosure?

4.1 Hypothesis Development

In Chapter 3, I document a robust positive association between AT and future voluntary forecasts. In 3.1, I discuss the evidence suggesting that managers change their voluntary disclosures in response to stock price informativeness, where more (less) informative prices lead to less (more) disclosure. The prior literature is mixed regarding the effect of AT on the informativeness of prices. AT appears to price known information more efficiently, which will more closely align prices with a manager's private information and result in less disclosure (Brogaard et al. 2014; Chakrabarty et al. 2016). There is some evidence, however, that AT can increase the cost of informed trading and reduce the total amount of information acquired (Hirschey 2016; Korajczyk and Murphy 2016; Weller 2016). In this case, AT would lead to less informative prices and more disclosure. The positive association between AT and disclosure that I find in Chapter 3 suggests the latter explanation may be more in line with the observed association.

Directly measuring the effect of AT on price informativeness is difficult because the amount of information in prices is not directly observable.[29] Instead, I develop predictions to test whether 1) AT is associated with a less informed market, 2) AT is associated with decreased information acquisition, 3) AT is associated with a higher cost of informed trading, and 4) the type of AT that might improve the informativeness of prices has the opposite effect on disclosure.

[29] As discussed previously, some of the prior literature tries to test this directly, finding varying results depending on the research design and windows over which price informativeness is measured.

First, I consider the market response to the forecast surprise. When the market updates prices in response to a firm disclosure, it considers the precision of its prior and the precision of the released news. When the precision of the market's prior is lower, the price response to the disclosure will be stronger (Verrecchia 1990). Rogers and Stocken (2005) document that investors place a lower precision on forecasts and respond less strongly when the forecast is expected to be more biased. If AT reduces the informativeness of prices, the market will place a lower precision on the current market price and thus I expect AT to be correlated with a higher management forecast response coefficient. This leads to my second hypothesis:

H2: AT is associated with a higher management forecast response coefficient.

Next, I test whether AT is associated with decreased information acquisition by investors. Because information acquisition is not observable directly, I perform three tests to try and capture this behavior. First, I utilize EDGAR downloads as a proxy for information acquisition. EDGAR downloads have previously been shown to improve market efficiency with respect to earnings announcements (Drake et al. 2014). Although these filings are publicly available, my assumption is that EDGAR downloads are correlated with other information acquisition activities and may still provide complementary information to an informed trader. As such, I state my third hypothesis as follows:

H3: AT is associated with fewer EDGAR downloads.

As an alternative proxy for information acquisition, I use analyst forecast revisions. Analysts have been shown to provide value-relevant information to the market (Beyer et al. 2010). One way analysts profit is by obtaining new information and providing it to their institutional clients in order to generate trading volume for their brokerage. As such, I assume that both analyst and investor objective functions overlap with respect to their desire to acquire private information. If AT reduces analyst information acquisition, I expect forecast revisions to be smaller in magnitude over the course of a quarter.

H4: AT is associated with a smaller magnitude of analyst forecast revisions.

Last, I combine analyst forecasts and market returns to proxy for the quantity of information acquired by the market over the course of a quarter. I make two assumptions in this test: 1) the market returns to an earnings announcement represent the true surprise to the announcement, and 2) the analyst consensus forecast at the beginning of the quarter is aligned with the market's consensus about upcoming earnings. I utilize the analyst consensus forecast, measured at the beginning of the quarter, as a measure of a stale earnings forecast. I predict that if the market acquires substantial new information over the course of the quarter, then the earnings surprise calculated from this stale consensus should weakly correlated with the returns to earnings (i.e., the true surprise). If the market does not acquire new information, however, then the stale consensus forecast will be more strongly correlated with the returns at the announcement. If AT reduces the amount of information acquired, the stale consensus should have a stronger response coefficient with market returns at the announcement.

H5: Stale earnings expectations are more strongly associated with the market returns to an earnings announcement when AT is high.

As I discuss in Chapter 2.2, AT is made up of a diverse set of trading strategies, some of which may have differing effects on price informativeness. My primary tests examine the net effect of these strategies on disclosure, finding a positive association. Some types of AT, however, may increase the informativeness of prices. When these types of AT play a larger role, they may mitigate the positive association between AT and disclosure.

Prior literature has documented that AT more efficiently prices information such as earnings, 8-K filings, news, and other hard information (Zhang 2012; Scholtus et al. 2014; Chakrabarty et al. 2016; Rogers et al. 2016). My assumption is that for a given level of AT, when the level of hard news released is high, a greater proportion of that AT is of the type that more efficiently prices the released news. I expect this type of AT to improve the informativeness of market prices, and therefore the reduce the level of disclosure.

H6: The effect of AT on disclosure is mitigated when the level of hard news during the quarter is high.

Chakrabarty et al. (2016) document that AT more efficiently prices earnings. If this is the case, then I expect a high level of AT during the prior earnings announcement to increase the informativeness of the market price and reduce disclosure during the following earnings announcement.

H7: AT during the prior quarter's earnings announcement is negatively correlated with the current quarter's forecasts.

The above hypotheses are based on the assumption that AT increases the cost of informed trading. If it is more expensive to trade on private information, traders are less likely to incur costly information acquisition activities. Therefore, I predict AT is associated with higher trading costs.

H8: AT is associated with a higher cost of informed trading.

Next, I try to understand more clearly what type of AT is driving my result. Although measuring specific AT strategies is not feasible, certain strategies (statistical arbitrage, event arbitrage) likely use active orders in their implementation, while others (market making, order execution) primarily use passive orders.[30] Prior literature finds active AT is more strongly associated with changes in price efficiency and information acquisition as compared to passive AT (Brogaard et al. 2014; Weller 2016). As such, I expect active strategies to be strongly correlated with changes in forecasts.

H9: Active AT strategies are more strongly correlated with management disclosure compared to passive AT strategies.

4.2 Data and Research Design

4.2.1 Forecast Response Coefficients

To test *H2*, I begin with the sample of all EPS forecasts over my sample period and calculate the forecast surprise (*ForecastSurprise*) as the manager's forecast less the analyst consensus at the time of the forecast, scaled by the analyst consensus. I measure the informativeness of the forecast as the abnormal returns to the forecast (*CAR_Forecast*). For each forecast, I measure the level of AT in the 30 days leading up to the forecast and interact this with the forecast surprise. I include control variables measured over the prior 30 days for firm size, analyst following, and various market measures.

[30] Order anticipation may be executed actively or passively.

$$
\begin{aligned}
CAR\_Forecast_{i,t} \\
&= \alpha + \beta_1 AT_{i,t} + \beta_2 ForecastSurprise_{i,t} + \beta_3 AT_{i,t} \\
&* ForecastSurprise_{i,t} + \beta_4 LnMktCap_{i,t} + \beta_5 AnalystFollowing_{i,t} \\
&+ \beta_6 CAR\_Pre_{i,t} + \beta_7 Turnover_{i,t} + \beta_8 Volatility_{i,t} + \beta TimeFE + \varepsilon_{i,t}
\end{aligned} \tag{5}
$$

The coefficient of interest is β_3 on the interaction of *AT* and *ForecastSurprise*. A positive (negative) coefficient shows the market reacts more (less) strongly to a given management forecast surprise when AT is high. This would suggest that the market is less (more) informed when AT leading up to a management forecast is high. I predict a positive coefficient on β_3.

4.2.2 Information Acquisition

As discussed in 4.1, I perform three tests to measure the effects of AT on information acquisition. For *H3*, I utilize EDGAR downloads. I begin with the sample discussed in Chapter 3.2.1. I obtain the EDGAR server logs for this sample and filter out requests from web crawlers, index page requests, and those with server codes of 300 or greater (these mostly return errors or messages indicating the file has been moved).[31] Following the prior literature, I classify downloads made from IP addresses that make 50 or less requests in a given day as non-robot, and downloads made from IP address that make more than 50 requests per day as robot downloads (Loughran and McDonald 2016). *NonRobotDownloads* and *RobotDownloads* capture the number of requests made by non-robots and robots, respectively, during the preannouncement period (i.e.,

[31] The sample period for this test runs from February 2012 through June 2015, due to the availability of the EDGAR server logs.

filing downloads measured contemporaneously with AT). I test the association between AT and SEC filing downloads in the following model:

$$
\begin{aligned}
Downloads_{i,t} = {} & \alpha + \beta_1 AT_{i,t} + \beta_2 NewFilings_{i,t} + \beta_3 LagUE_{i,t} + \beta_4 LagCAR\_Ann_{i,t} \\
& + \beta_5 LagLoss_{i,t} + \beta_6 LagSalesGrowth + \beta_7 LagLeverage_{i,t} \\
& + \beta_8 LagAccruals_{i,t} + \beta_9 LagDividends + \beta_{10} GuideCount\_Pre_{i,t} \\
& + \beta_{11} GuideCount\_Pre2_{i,t} + \beta_{12} CAR\_Pre_{i,t} + \beta_{13} InsiderSales_{i,t} \\
& + \beta_{14} NewsVolume_{i,t} + \beta_{15} PPUE_{i,t} + \beta_{16} LnMktCap_{i,t} \\
& + \beta_{17} AnalystFollowing_{i,t} + \beta_{18} Dispersion_{i,t} + \beta_{19} Turnover_{i,t} \\
& + \beta_{20} Volatility_{i,t} + \beta_{21} Skewness_{i,t} + \beta_{22} Spread_{i,t} + \beta TimeFE \\
& + \beta FirmFE + \varepsilon_{i,t}
\end{aligned}
\tag{6}
$$

β_1 is my coefficient of interest; I expect it to be negative if AT is associated with decreased information acquisition. I include a control for the number of new SEC filings (*NewFilings*) made during the prior earnings announcement window and throughout the current preannouncement window. I also control for information released during the prior earnings announcement and current quarter, along with time and firm fixed effects (see Appendix A for variable definitions; lagged variables are measured as of the prior quarter).

For *H4*, my second proxy for information acquisition is *AbsFrcstChange*, calculated as the absolute value of the consensus analyst forecast as of the beginning of the preannouncement period less the consensus analyst forecast at the end of the preannouncement period, scaled by price. I assume that a greater magnitude change in forecasts indicates analysts have acquired more information over the quarter leading up to the earnings announcement. To test this, I utilize equation (6) with the following adjustments. I include the absolute value of unexpected earnings calculated using the consensus forecast at the beginning of the preannouncement period (*AbsUE_BegQtr*). This variable controls for the magnitude of the information analysts have the opportunity to acquire over the quarter. I add the standard deviation of *AbsFrcstChange* for the firm over the sample period (*FrcstChgStd*), and measure the number of analysts and forecast dispersion as of the beginning of the quarter. I remove firm fixed effects as I expect *AbsFrcstChange* to vary with the analyst error at the beginning of the quarter (*AbsUE_BegQtr*), rather than a firm level effect.

Last, I consider whether the reduced information acquisition manifests itself in the market returns to earnings announcements. To test *H5*, I obtain the cumulative abnormal returns during the earnings announcement window (*CAR_Ann*) and consider this to be a proxy for the level of surprise investors have to earnings news. Next, I measure the consensus forecast at the beginning of the preannouncement period (approximately a quarter prior the earnings announcement), and use this expectation to calculate unexpected earnings (*UE_BegQtr*). I propose that if the earnings surprise using quarter old expectations, *UE_BegQtr*, is highly correlated with the market surprise at the time of the announcement, *CAR_Ann*, then investors acquired little information during the preannouncement period. If, however, *UE_BegQtr* is uncorrelated with the market surprise, then investors fully updated their expectations during the quarter. As my dependent variable is announcement period abnormal returns, I include accounting metrics as controls:

$$\begin{aligned} CAR\_Ann_{i,t} = {} & \alpha + \beta_1 AT_{i,t} + \beta_2 UE\_BegQtr_{i,t} + \beta_3 AT_{i,t} * UE\_BegQtr_{i,t} \\ & + \beta_4 GuideCount\_Ann_{i,t} + \beta_5 GuideCount\_Pre_{i,t} \\ & + \beta_6 GuideCount\_Pre2_{i,t} + \beta_7 Loss_{i,t} + \beta_8 SalesChange_{i,t} \\ & + \beta_9 CAR\_Pre_{i,t} + \beta_{10} PPUE_{i,t} + \beta_{11} LnMktCap_{i,t} \\ & + \beta_{12} AnalystFollowing_{i,t} + \beta_{13} Dispersion_{i,t} + \beta_{14} Turnover_{i,t} \\ & + \beta_{15} Volatility_{i,t} + \beta_{16} Skewness_{i,t} + \beta_{17} Spread_{i,t} + \beta_{18} InsiderSales \\ & + \beta_{19} NewsVolume_{i,t} + \beta_{20} Leverage_{i,t} + \beta_{21} Accruals_{i,t} \\ & + \beta_{22} Dividends_{i,t} + \beta TimeFE + \varepsilon_{i,t} \end{aligned} \tag{7}$$

See Appendix A for variable definitions. My coefficient of interest is β_3, the interaction of *AT* and *UE_BegQtr*. A positive (negative) coefficient suggests investors acquire less (more) information when AT is high. I predict a positive coefficient on β_3.

4.2.3 Informativeness Improving AT

In *H6* and *H7*, I predict that event arbitrage AT has a negative effect on future forecasts. To test *H6*, I begin with my main sample discussed in Chapter 3.2.1 and run the following model:

$$\begin{aligned} Guider_{i,t} = {} & \alpha + \beta_1 AT_{i,t} + \beta_2 NewsVolume_{i,t} + \beta_3 AT_{i,t} * NewsVolume_{i,t} + \beta Controls \\ & + \beta TimeFE + \beta FirmFE + \varepsilon_{i,t} \end{aligned} \tag{8}$$

The coefficient of interest is β_3, the interaction of AT and *NewsVolume*. I predict a negative coefficient on β_3, which would suggest that event arbitrage AT has an opposite effect on disclosure. The controls are the same as in equation (1).

To test *H7*, I measure the level of AT during the prior quarter's earnings announcement window (*LagAT_Announcement*) and run the following logit model.

$$Guider_{i,t} = \alpha + \beta_1 AT_{i,t} + \beta_2 LagAT\_Announcement_{i,t} + \beta Controls + \beta TimeFE + \beta FirmFE + \varepsilon_{i,t} \tag{9}$$

I predict a negative coefficient on β_2. This would suggest that the type of AT that prices earnings more efficiently and makes prices more informative has the opposite effect on future forecasts.

4.2.4 Cost of Informed Trading

I predict in *H8* that AT is associated with a higher cost of informed trading. I proxy for the cost of informed trading using the Amihud (2002) illiquidity measure (*Illiquidity*), calculated as the absolute value of stock returns divided by the dollar volume traded measured daily and averaged over the pre-announcement period. A higher ratio indicates returns move more for a given trade, which increases the cost of trading on private information. I run the following model over the pre-announcement period, controlling for firm characteristics and contemporaneous market measures:

$$Illiquidity_{i,t} = \alpha + \beta_1 AT_{i,t} + \beta_2 LnMktCap_{i,t} + \beta_2 PPUE_{i,t} + \beta_3 InsiderSales_{i,t} + \beta_4 AnalystFollowing_{i,t} + \beta_5 Dispersion_{i,t} + \beta_6 Spread_{i,t} + \beta_7 Turnover_{i,t} + \beta_8 Volatility_{i,t} + \beta_9 Skewness_{i,t} + \beta_{10} NewsVolume_{i,t} + \beta TimeFE + \beta FirmFE + \varepsilon_{i,t} \quad (10)$$

I predict a positive coefficient on β_1, which would suggest AT increases the cost of informed trading.

4.2.5 Active vs. Passive AT

In *H9*, I predict active AT strategies have a stronger effect on disclosure than passive AT strategies. Of my AT proxies, *OddLotVolumeRatio* and *TradeSize* are associated with active AT, as the recording of an odd lot trade or trade size is more likely dictated by the trade that crosses the spread (Weller 2016).[32] Likewise, the *CancellationToTradeRatio* and *TradeToOrderRatio* are more likely dictated by passive AT which generate large numbers of limit orders.

I create two proxies for active and passive AT. *LiquidityMakingAT* is a principal components analysis of *CancelToTradeRatio* and *TradeToOrderRatio*. The *CancelToTradeRatio* (*TradeToOrderRatio*) is positively (negatively) associated with *LiquidityMakingAT*, and therefore I expect *LiquidityMakingAT* to be positively associated with guidance. *LiquidityTakingAT* is a principal components analysis of *OddLotVolumeRatio* and *TradeSize*, where the former is positively correlated with *LiquidityTakingAT* and the latter is negatively correlated with

[32] On some exchanges, the passive side of the trade dictates the trade size if it is smaller than the active order, resulting in the active order executing against multiple passive orders. Weller (2016) provides evidence that the method of reporting trade sizes does not materially affect trade size and the odd-lot volume ratio, and therefore they are appropriate proxies for active AT. Likewise, his evidence suggests that the cancellation to trade ratio and trade to order ratio primarily vary based on the number of passive orders placed, not the level of active trading.

LiquidityTakingAT. Therefore, I expect *LiquidityTakingAT* to be positively correlated with guidance. For each principal components analysis, the first principal component has an eigenvalue of 1.8 and explains 88% of the variance. The correlations between the proxies are consistent with expectations as well: Table 2 shows the *OddLotVolumeRatio* and *TradeSize* are strongly negatively correlated, and the *TradeToOrderRatio* and *CancelToTradeRatio* are negatively correlated as well.

I use my main sample from Chapter 3.2.1 and equation (1), but replace *AT* with the active and passive AT proxies discussed above. In addition, I regress guidance on the individual AT proxies as well.

4.3 Results

In Table 6, I present my results for *H2*. The coefficient of interest is the interaction between *AT* and *ForecastSurprise*. The positive coefficient (significant at 5%) that for a given level of forecast news, the market reaction is stronger when AT leading up to the forecast is high. This is consistent with the market knowing less and therefore placing a stronger weight on the management forecast when updating the price.[33]

Table 7, Panel A presents my results for *H3*. In columns (1) and (2), I find AT is negatively correlated with *NonRobotDownloads* (significant at 1%), but is not significantly correlated with *RobotDownloads.* Due to the risk of reverse causality with the contemporaneous measurement of AT and downloads, I also measure SEC downloads in the quarter strictly following my measurement of AT (see columns (3) and (4) for *LeadNonRobotDownloads* and

[33] The result is robust to controlling for the announcement window AT. In untabulated tests, a similar result holds for earnings surprises; the earnings response coefficient is higher when AT leading up to the announcement is high.

LeadRobotDownloads, respectively).[34] I find similar results; AT is negatively associated with non-robot downloads in the following quarter (significant at 5%), but is not significantly correlated with robot downloads. These results are consistent with AT reducing information acquisition by non-algorithmic traders.

Panel B of Table 7 shows *AT* is negatively associated with *AbsFrcstChange* (significant at the 1% level) in column (1), as predicted in *H4*. This result implies analysts update their forecasts less when AT is high and therefore acquire less information. Because AT and the forecast change are measured over the same period, there may be concerns of reverse causality such that the smaller analyst forecast revisions cause greater AT. To account for this, I measure the absolute analyst forecast change in the quarter following my AT measurement (*LeadAbsFrcstChange*).[35] The result is similar as seen in column (2) of Panel B (significant at the 5% level).

I present the results of *H5* in Table 7, Panel C. Column (1) shows a positive coefficient for the interaction between *AT* and *UE_BegQtr* (significant at 1%). This finding indicates that outdated earnings expectations, *UE_BegQtr*, are more highly correlated with the true announcement surprise (abnormal returns) when AT is high. This result suggests that investors acquire less information in the quarter leading up to an earnings announcement with greater AT. Together, these three tests provide evidence that AT discourages information acquisition, leaving firms to replace the lost information with their own disclosures.

[34] I also update the controls to measure new SEC filings in the same quarter in which I measure downloads, and control for the current earnings announcement information and returns, instead of lagged variables.

[35] Similar to my test of SEC filing downloads, I update the controls to the values known at the beginning of the quarter over which *LeadAbsFrcstChange* is measured (e.g., instead of *AbsUEBegQtr*, I include *LeadAbsUE_BegQtr*).

Table 8 presents the results for *H6* and *H7* on informativeness improving AT. In column (1), I interact *AT* with the level of *NewsVolume* over the *AT* measurement period. The coefficient on the interaction term is negative and significant at the 10% level. This result suggests that when news volume is high, the positive association between AT and disclosure is mitigated. This is consistent with event arbitrage AT pricing the released news more efficiently, leading to greater price informativeness.

In Table 8, column (2), I include the prior quarter's announcement window AT (*LagAT_Announcement*). The coefficient on *LagAT_Announcement* is negative and significant at 5%, consistent with expectations. Again, this suggests that informativeness improving AT has a negative effect on future management forecasts. These results are interesting because they show that different types of AT can have opposite effects on disclosure.

I test *H8* in Table 9, finding a positive and significant coefficient on *AT*. This result suggests that when AT is high, the absolute price movement for a given level of trading volume is higher. When illiquidity is higher, trading on private information is more expensive. This result is consistent with AT increasing the cost of informed trading, which may discourage investors from becoming informed.

Last, I test *H9* in Table 10. In Panel A, I regress the guidance indicator, *Guider*, on the active and passive AT proxies. Individually, the proxies are both positive and significant at 1% (columns (1) and (2) in Table 10 Panel A). When placed in the same regression, *LiquidityTakingAT* is significant at 5%, while *LiquidityMakingAT* remains significant at 1%. The coefficients on the two variables, however, are nearly identical. In contrast to the prior literature which finds stronger results for liquidity taking AT, the results in Table 10 suggest that both active and passive AT play a role in increased disclosure, indicating market making AT's ability to avoid adverse selection may increase the cost of informed trading and lead to less precise market prices.

In Panel B of Table 10, I test the individual AT proxies. I find the active AT proxies, *OddLotVolumeRatio* and *TradeSize*, are significantly associated with *Guider* in the expected direction at 5% and 1%, respectively.[36] Similarly, the passive AT proxies, *CancelToTradeRatio* and the *TradeToOrderRatio*, are significantly associated with *Guider* at 1%.

4.4 Conclusion

In this chapter, I examine why I find a positive association between AT and disclosure in Chapter 3. I predict that this positive association may be the result of less informative market prices due to investors acquiring less information. I find the market reacts more strongly to forecasts when AT is high, suggesting the precision of the market's prior is lower and the market may be less informed. Moreover, I find evidence that AT is associated with decreased information acquisition by investors and analysts. AT is associated with a higher cost of informed trading, which may explain the decreased information acquisition. I also find that informativeness improving AT has the opposite effect on voluntary disclosures. As a whole, the data is consistent with the explanation that AT reduces the informativeness of prices by discouraging investors from becoming informed, leaving managers to inform the market through voluntary disclosures instead.

---

[36] Recall the *OddLotVolumeRatio* and *CancelToTradeRatio* are positively correlated with AT, while the *TradeToOrderRatio* and *TradeSize* are negatively correlated with AT. Given the positive association between AT and guidance in Table 3, I expect the former two individual proxies to be positively associated with guidance, and the latter two to be negatively associated with guidance.

CHAPTER 5

Algorithmic Trading and Other Disclosure Characteristics

5.1 Hypotheses Development

In this chapter, I examine how AT affects other disclosure characteristics beyond the quantity of forecasts disclosed. If AT decreases price informativeness, managers have the option to increase the precision of their disclosures in addition to, or instead of, disclosing more. A greater precision should induce a stronger market reaction as well (Verrecchia 1990). Li and Zhang (2015) show that managers respond to increased short-selling pressure by reducing the precision of their disclosures. As such, I predict at AT is positively associated with forecast precision:

H10: AT is positively associated with forecast precision.

Next, I consider whether AT affects the timeliness of disclosures. Managers face a tradeoff between disclosing the information themselves and incurring the cost of disclosure, or waiting to see if investors acquire the information on their own. If AT decreases information acquisition, I predict managers disclose in a more timely manner, e.g., forecast earlier and with a greater horizon, as there is a lower likelihood the information will be acquired and priced by investors.

H11: AT is positively associated with forecast timeliness.

Last, I consider whether AT affects the firm's soft information disclosures. Prior research shows that AT is adept at incorporating hard, but not soft, information into stock prices (Zhang 2012). If so, managers may improve their soft information disclosures when AT is high. I consider an improvement (reduction) in readability to be an increase (decrease) in the quality of soft disclosures.[37]

H12: AT is positively associated with the readability of firms' annual and quarterly reports.

5.2 Data and Research Design

First, I measure the specificity of the firm's quantitative disclosures to test *H10*. I extract the EPS forecast with the longest horizon for each firm-quarter from my main sample described in Chapter 3.2.1. Using this forecast, I create the variable *Specificity,* which equals 3 if the forecast is a point estimate, 2 if the forecast is a closed range (i.e., the manager provides an upper and lower bound), and 1 if the forecast is an open range (i.e., the manager provides one numeric value and indicates EPS is expected to be above or below that value).

My model is the same as equation (1), but using *Specificity* as the dependent variable and four additional controls: *GuideCount_Pre, GuideCount*, *Annual*, and *Horizon_EPS*. *GuideCount_Pre* is the count of the number of pieces of guidance issued in the pre-announcement period. *GuideCount* is the count of guidance in the announcement window. *Horizon_EPS* is the horizon of the forecast being examined. *Annual* is an indicator equal to 1 if it is the firm's 4th fiscal quarter, and 0 otherwise. I include these controls because specificity of forecasts may vary with the number of forecasts they disclose, whether it is an annual or quarterly report, and the horizon of the forecast.

[37] I recognize this assumes the content of the report is held constant. Some may argue an improvement in readability could be associated with a loss of detail or technical nuance in the report. Nevertheless, I believe testing the effect of AT on readability to be valuable.

To test *H11*, I develop a range of proxies for both mandatory and voluntary disclosure timeliness. I test five measures: the number of days between the quarter end and the earnings announcement date (*EADelay*), the number of days between the earnings announcement date and the 10-K or 10-Q filing date (*Filing_Delay*), the horizon of the longest horizon EPS forecast (*Horizon_EPS*), an indicator for whether the firm's EPS forecast is bundled with the earnings announcement (*Bundle_EPS*), and an indicator for whether the firm announces outside of regular trading hours (*PostClose*). I use my main model from equation (1) with firm fixed effects.

Last, for *H12*, I measure *Readability* as the first principal component of the eight readability measures provided as part of the WRDS SEC Analytics Suite.[38] A high value of *Readability* is associated with more readable reports. My model for *H12* is as follows:

$$\begin{aligned} Readability_{i,t} &= \alpha + \beta_1 AT_{i,t} + \beta_2 UE_{i,t} + \beta_3 PosUE_{i,t} + \beta_4 NegUE_{i,t} + \beta_5 Loss_{i,t} \\ &+ \beta_6 CAR\_Pre_{i,t} + \beta_7 PPUE_{i,t} + \beta_8 LnMktCap_{i,t} \\ &+ \beta_9 AnalystFollowing_{i,t} + \beta_{10} Dispersion_{i,t} + \beta_{11} SalesGrowth_{i,t} \\ &+ \beta_{12} Turnover_{i,t} + \beta_{13} Volatility_{i,t} + \beta_{14} Skewness_{i,t} + \beta_{15} Spread_{i,t} \\ &+ \beta_{16} InsiderSales_{i,t} + \beta_{17} Annual_{i,t} + \beta_{18} GuideCount\_Pre_{i,t} \\ &+ \beta_{19} GuideCount_{i,t} + \beta FirmFE + \beta TimeFE + \varepsilon_{i,t} \end{aligned} \tag{9}$$

[38] The eight readability measures are the Flesch Reading Ease Index, Flesch-Kincaid Readability Index, RIX Readability Index, Coleman Readability Index, Gunning Fog Readability Index, Automated Readability Index, Smog Readability Index, and LIX Readability Index. These measures were developed for uses outside of accounting and finance but have been adopted, especially the Gunning Fog Readability Index, frequently over the last 10 years (Li 2008, Loughran and McDonald 2014). The eight measures are correlated as expected, and the first principal component has an eigenvalue of 6.73.

I add the variable *Annual*, which is equal to 1 if it is the firm's fiscal fourth quarter, as well as measures for the count of guidance issued during the disclosure period (*GuideCount*) and preannouncement window (*GuideCount_Pre*).

Loughran and McDonald (2014) argue the file size of the 10-K or 10-Q is a better proxy for readability, where smaller files are more readable. As such, I obtain the associated 10-K or 10-Q file size from the WRDS SEC Analytics Suite and scale it by 1,000,000 to generate *FileSize*.

5.3 Results

I present the results for *H10* in Table 11. In column (1), I find *AT* is negatively correlated with *Specificity* (significant at 5%), indicating firms are more likely to forecast a range as opposed to a point estimate, and the range is more likely to be an open range than a closed range, when AT is high.

For firms that issue a closed range EPS forecast, I measure the magnitude of the forecast range (*AbsFrcstRange*) as the absolute value of the difference between the upper and lower bounds forecasted, scaled by price. As shown in Table 11, column (2), *AT* is negatively associated with *AbsFrcstRange* (significant at the 5% level). This indicates ranged forecasts become more precise with greater AT, which is in contrast to the results for *Specificity*.[39] It is not clear why we would find differing results for these two proxies of forecast precision.

Table 12 presents the results for *H11* on forecast timeliness. I find AT is not significantly associated with any of the timeliness measures. It may be that other factors play a much larger role in determining the earnings announcement and filing dates, such as when the audit is

[39] This result is consistent if I restrict the sample to firms that issued a closed range forecast in the prior quarter, suggesting it is not due to firms previously issuing point estimates switching to closed range forecasts.

completed. Regarding *PostClose*, announcement times are often very persistent and therefore AT may not be a significant enough factor.

In Table 13, I present the results for *H12* on readability. The coefficients on AT for both *Readability* and *FileSize* are positive but not significant. This may be the result of using the annual and quarterly reports to measure readability. These mandated reports are the most formal and least flexible of firm written disclosures, and therefore detecting a change may be difficult.

5.4 Conclusion

In Chapter 5, I test the association between AT and other disclosure characteristics. Managers have the opportunity to not only vary the quantity of forecasts they make, but also the precision, timeliness, and readability of soft disclosures. I find mixed results: the type of forecast becomes less specific when AT is high, but a closed range forecast becomes more specific. I find no change in the timeliness of mandatory and voluntary disclosures. The results regarding the readability of annual and quarterly reports are both positive, but not significant. An alternative, and potentially superior, test may be to investigate the readability of less formal disclosures such as press releases or conference calls.

CHAPTER 6

Conclusion

I investigate the effects of AT on managerial disclosure decisions. I find AT is positively associated with both the likelihood of issuing guidance and the quantity of guidance issued. My results are robust to the inclusion of firm fixed effects, prior guidance, and controls for information events during the AT measurement period. A natural experiment using the implementation of the NYSE Autoquote as an exogenous shock to AT supports this finding, as does inverse-propensity reweighting techniques.

I find evidence that forecasts are more informative and traders acquire less information when AT is high. These results are consistent with AT reducing the incentives to become informed, leading to a less informed market and greater need for disclosure. These effects are mitigated when the AT is more likely to be pricing earnings information or other news. Moreover, I find some evidence that informed trading becomes more expensive when AT is present, which could explain the decrease in information acquisition. Both active and passive AT is associated with increased disclosure, contrary to expectations. Further, I find mixed results on the effect of AT on forecast specificity, and no association between AT and the timing of disclosures.

Prior research has focused on the effects of AT on market outcomes such as liquidity and price efficiency. I contribute to the literature by being the first to provide evidence that AT affects the decision making of managers. Managers play an integral role in shaping the information environment of the firm and markets, and therefore it is critical to understand how they respond to changes in market microstructure.

I also contribute to the literature by identifying a new pathway through which AT affects the information environment and information asymmetry between firms and investors. The prior finance literature finds evidence that AT improves the pricing of known information, but may deter investors from becoming informed. I find that managers respond as well, increasing the quantity of guidance issued.

Areas for future research could include more thoroughly investigating how AT affects firms' soft disclosures as well as how AT affects the pricing of soft disclosures. Given that AT is the first type of trading to price firm news, we may also see changes in firms' propensity to manage earnings or the timing of insider trades. If AT is more (less) adept than human traders at detecting earnings management and therefore decreases (increases) the benefit of managing earnings, we may observe less (more) earnings management. Likewise, if AT accelerates the pricing of firm news such that post-announcement trades become less profitable, we may observe an increasing number of management trades outside of approved windows.

FIGURES

Figure 1

Figure 1 presents the average *GuideCountEPS* per quarter of the matched Autoquote sample for treated and untreated firms from June 2000 to June 2006.

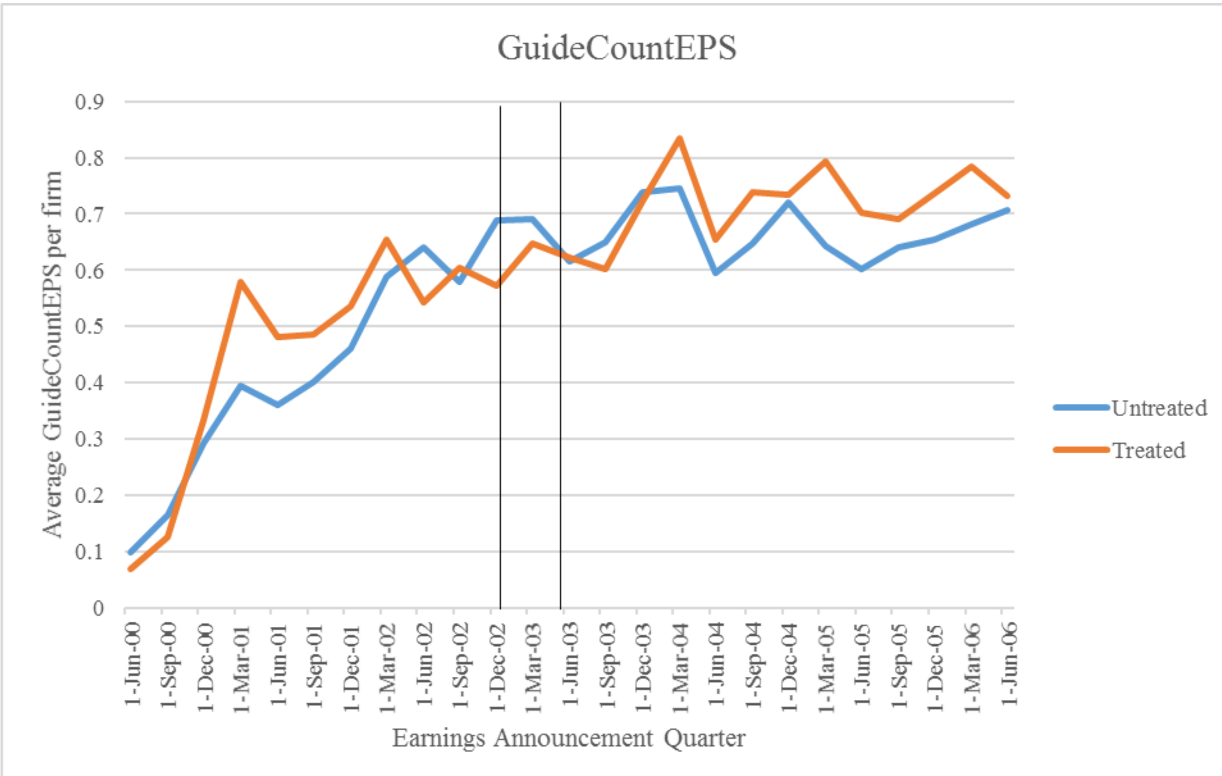


Figure 2

Figure 2 presents the average *GuideCount* per quarter of the matched Autoquote sample for treated and untreated firms from June 2000 to June 2006.

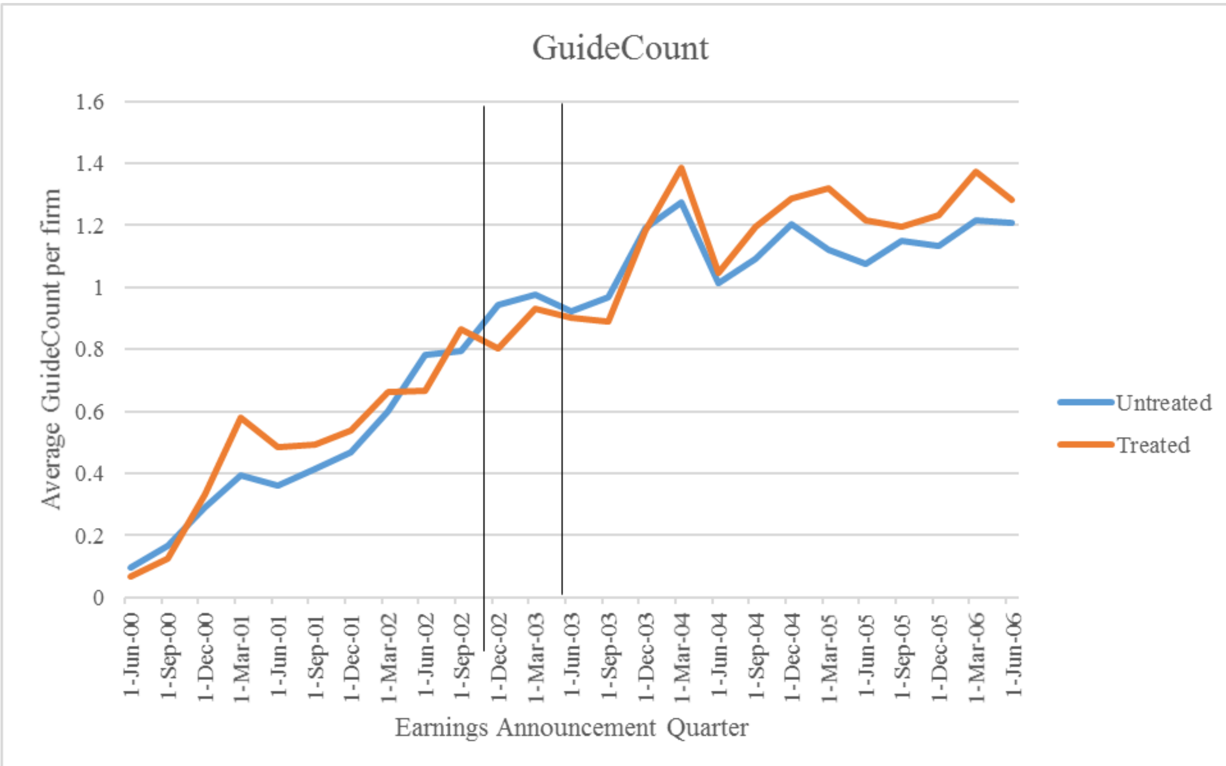


Figure 3

Figure 3 presents the kernel density estimate to be in AT Quintile 1 for each observed AT Quintile before trimming. AT Quintile 5 represents firms with the highest observed AT, and AT Quintile 1 represents firms with the lowest observed AT.

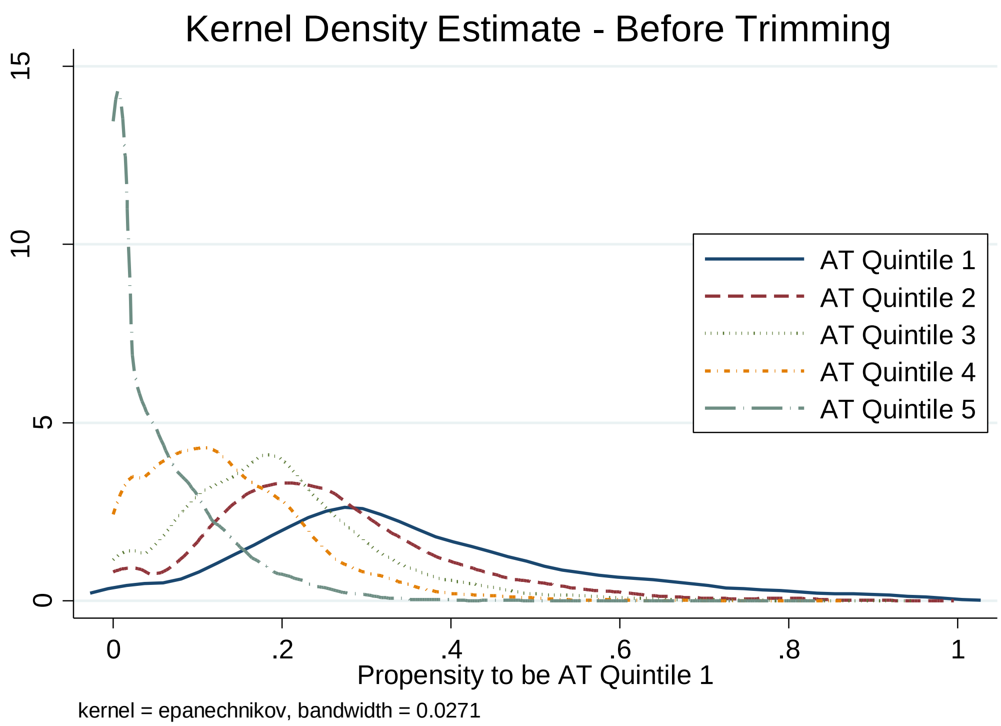
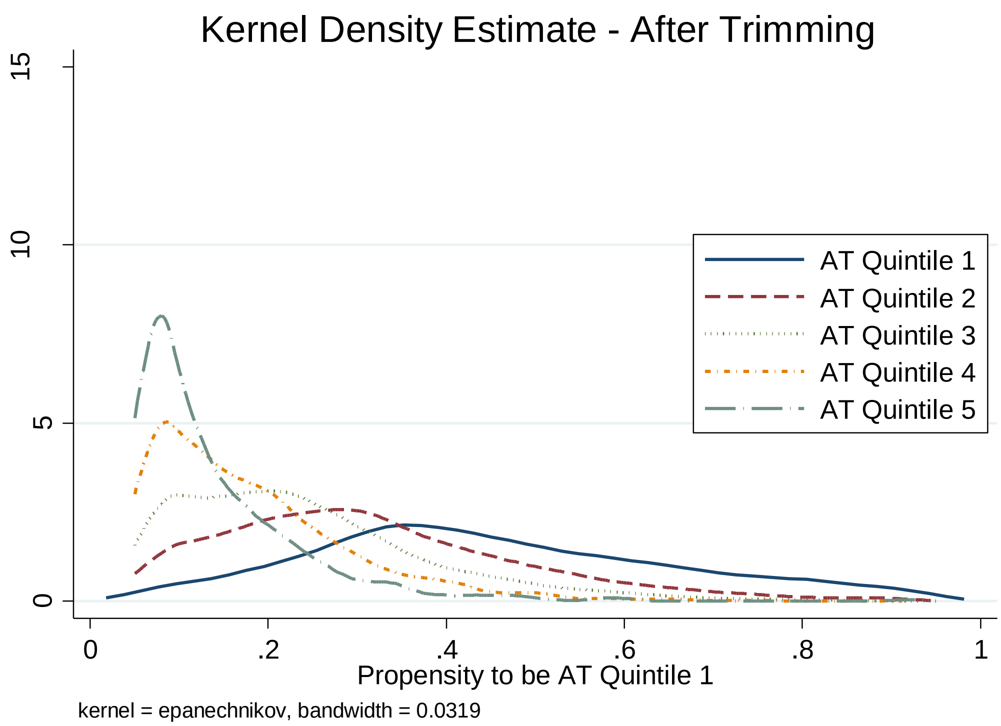


Figure 4

Figure 4 presents the kernel density estimate to be in AT Quintile 1 for each observed AT Quintile after trimming. The sample was trimmed at 0.05 and 0.95 three times for the likelihood of being AT Quintile 1. AT Quintile 5 represents firms with the highest observed AT, and AT Quintile 1 represents firms with the lowest observed AT.



TABLES

Table 1
Summary Statistics (N=34,310)

Table 1 presents summary statistics for variables used in the empirical tests. See Appendix A for variable definitions.

Variable	Mean	StdDev	P25	Median	P75
AT Proxies					
OddLotVolumeRatio	-2.150	0.557	-2.455	-2.086	-1.751
TradeToOrderRatio	-3.783	0.523	-4.110	-3.744	-3.410
CancelToTradeRatio	3.397	0.522	3.056	3.328	3.664
TradeSize	4.770	0.308	4.547	4.763	4.969
AT	-0.061	1.516	-0.993	-0.026	0.925
Disclosure Variables					
Guider	0.712	0.453	0.000	1.000	1.000
GuideCount	2.223	2.293	0.000	2.000	3.000
GuideDays	1.043	0.901	0.000	1.000	1.000
Specificity_EPS	2.049	0.351	2.000	2.000	2.000
AbsFrcstRange_EPS	0.004	0.010	0.001	0.003	0.005
EADelay	34.810	11.900	27.000	33.000	39.000
FilingDelay	132.472	190.392	1.000	14.000	247.000
Horizon_EPS	238.674	178.143	92.000	184.000	365.000
Bundle_EPS	0.893	0.309	1.000	1.000	1.000
PostClose	0.922	0.269	1.000	1.000	1.000

Table 1, *continued*
Summary Statistics (N=34,310)

Control Variables					
UE	-0.001	0.041	-0.001	0.000	0.002
PosUE	0.579	0.494	0.000	1.000	1.000
NegUE	0.359	0.480	0.000	0.000	1.000
PPUE	0.616	0.310	0.500	0.750	1.000
Loss	0.229	0.420	0.000	0.000	0.000
SalesGrowth	0.306	12.125	-0.025	0.055	0.163
Annual	0.266	0.442	0.000	0.000	1.000
CAR_Pre	-0.025	0.210	-0.102	-0.006	0.080
LogMktCap	14.442	1.603	13.277	14.309	15.444
AnalystFollowing	12.247	8.669	6.000	10.000	17.000
Dispersion	0.071	0.229	0.018	0.035	0.070
Spread	0.028	0.073	0.011	0.014	0.024
Turnover	9.877	9.445	5.001	7.615	11.863
Volatility	0.023	0.011	0.015	0.021	0.028
Skewness	0.120	1.450	-0.400	0.095	0.636
InsiderSales	1.809	26.068	0.000	0.000	0.196
NewsVolume	72.604	258.625	13.000	28.000	65.000

Table 2
Correlation Matrix

Table 2 presents the correlation matrix for the four individual AT proxies along with the final AT variable. See Appendix A for variable definitions. * indicates statistical significance at 10%.

	Variable	(1)	(2)	(3)	(4)	(5)
(1)	*OddLotVolumeRatio*	1				
(2)	*TradeToOrderRatio*	-0.59*	1			
(3)	*CancelToTradeRatio*	0.57*	-0.78*	1		
(4)	*TradeSize*	-0.76*	0.22*	-0.30*	1	
(5)	*AT*	0.91*	-0.81*	0.82*	-0.69*	1

TABLE 3
AT on Guidance Likelihood and Frequency

Table 3 presents the results of regressing guidance on *AT* . The dependent variable is one of three variables, *Guider* , *GuideCount* , or *GuideDays*. Columns 1-3 of Panel A present the general OLS model. Columns 4-6 of Panel A present the firm fixed effects model. Panel B presents a first differences model. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Panel A: Base and Fixed Effect Models

	Base			Firm Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	*Guider*	*GuideCount*	*GuideDays*	*Guider*	*GuideCount*	*GuideDays*
AT	0.130***	0.062**	0.013	0.172***	0.046***	0.020***
	(4.44)	(2.25)	(1.30)	(3.87)	(3.28)	(3.06)
UE	-1.005**	-0.828***	-0.186*	0.468	0.206	0.010
	(-2.30)	(-2.59)	(-1.82)	(1.02)	(1.06)	(0.26)
PosUE	-0.063	-0.075	-0.001	0.237**	0.109***	0.018
	(-0.84)	(-1.07)	(-0.03)	(2.18)	(3.86)	(0.99)
NegUE	-0.348***	-0.509***	-0.090***	0.034	0.059**	-0.022
	(-4.53)	(-7.12)	(-3.31)	(0.31)	(2.00)	(-1.21)
PPUE	0.666***	1.259***	0.181***	0.183	0.071*	0.027
	(8.67)	(14.68)	(6.95)	(1.54)	(1.90)	(1.49)
Loss	-0.076	0.160**	-0.026	-0.040	-0.025	-0.013
	(-1.23)	(2.04)	(-1.27)	(-0.48)	(-0.91)	(-0.92)
CAR_Pre	-0.020	0.028	0.011	-0.071	-0.003	0.008
	(-0.35)	(0.56)	(0.50)	(-0.66)	(-0.08)	(0.43)
LnMktCap	0.016	-0.020	0.054***	0.235*	0.208***	0.101***
	(0.39)	(-0.59)	(4.36)	(1.82)	(4.72)	(5.18)
SalesGrowth	-0.045	-0.002***	-0.001***	0.008	0.000	-0.000
	(-1.16)	(-3.74)	(-4.00)	(0.54)	(0.60)	(-0.28)
InsiderSales	0.000	0.001	-0.000	0.001	0.001*	0.000
	(0.74)	(1.64)	(-0.43)	(1.11)	(1.94)	(0.71)
AnalystFollowing	0.045***	0.033***	0.017***	0.052***	0.010**	0.005*
	(6.17)	(5.44)	(7.54)	(3.12)	(2.06)	(1.91)
Dispersion	-0.829***	-0.517***	-0.108***	-0.023	-0.020	-0.004
	(-2.66)	(-2.74)	(-2.71)	(-0.16)	(-0.88)	(-0.43)
Spread	-9.639***	-2.295**	-1.085***	-5.965***	-0.335	-0.195
	(-5.08)	(-2.44)	(-2.84)	(-2.69)	(-1.63)	(-1.57)
Turnover	0.009	0.003	0.004**	0.009	0.000	0.003**
	(1.43)	(0.62)	(2.28)	(1.34)	(0.04)	(2.49)
Volatility	-17.282***	-8.704***	-6.848***	-5.658	-1.008	-2.461***
	(-4.37)	(-2.62)	(-5.74)	(-1.01)	(-0.61)	(-3.30)

TABLE 3, *continued*
AT on Guidance Likelihood and Frequency

Panel A regressions continued from prior page.

Skewness	-0.056***	-0.055***	-0.013***	-0.056***	-0.014**	-0.002
	(-4.93)	(-4.27)	(-3.14)	(-2.89)	(-2.24)	(-0.65)
NewsVolume	-0.000**	-0.000***	-0.000***	-0.001***	-0.000**	-0.000**
	(-2.36)	(-3.87)	(-5.13)	(-2.79)	(-2.11)	(-2.23)
Constant	1.126*	2.040***	0.259		-0.848	-0.380
	(1.88)	(4.03)	(1.43)		(-1.33)	(-1.37)
Firm Fixed Effects	N	N	N	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
R-squared	0.094	0.082	0.113	0.021	0.774	0.518
N	34,310	34,310	34,310	15,083	34,310	34,310

TABLE 3, *continued*
AT on Guidance Likelihood and Frequency

Panel B: First Differences

Variable	(1) *Guider_Change*	(2) *GuideCount_Change*	(3) *GuideDays_Change*
AT_Change	0.1797***	0.1297***	0.0477***
	(4.70)	(5.64)	(6.15)
UE_Change	-0.1114	-0.0539	-0.0211
	(-0.39)	(-0.35)	(-0.46)
PosUE_Change	-0.0090	0.0999*	0.0265
	(-0.12)	(1.82)	(1.38)
NegUE_Change	-0.0853	0.0427	-0.0112
	(-1.16)	(0.76)	(-0.57)
PPUE_Change	-0.0755	-0.0416	-0.0186
	(-0.64)	(-0.58)	(-0.68)
Loss_Change	0.0088	-0.0208	0.0052
	(0.14)	(-0.49)	(0.34)
CAR_Pre_Change	-0.0437	0.0844	0.0089
	(-0.54)	(1.59)	(0.48)
LnMktCap_Change	0.2201	0.2366***	0.0667**
	(1.51)	(2.78)	(2.34)
InsiderSales_Change	0.0002	0.0003	0.0001
	(0.76)	(0.89)	(0.95)
SalesGrowth_Change	0.0001	0.0001	0.0001
	(0.40)	(0.76)	(1.09)
AnalystFollowing_Change	0.0205	0.0163	-0.0013
	(1.35)	(1.53)	(-0.31)
Dispersion_Change	-0.0429	-0.0299	0.0011
	(-0.84)	(-0.79)	(0.08)
Spread_Change	-0.5657	-0.1117	-0.0653
	(-0.84)	(-0.40)	(-0.61)
Turnover_Change	-0.0008	0.0013	-0.0001
	(-0.16)	(0.42)	(-0.07)
Volatility_Change	-0.5525	-0.6717	-1.0528
	(-0.12)	(-0.28)	(-1.27)
Skewness_Change	-0.0285**	0.0015	0.0031
	(-1.97)	(0.14)	(0.87)
NewsVolume_Change	-0.0001	-0.0003**	-0.0001*
	(-0.88)	(-2.22)	(-1.90)

TABLE 3, *continued*
AT on Guidance Likelihood and Frequency

Panel B regressions continued from prior page.

Constant		-0.1611***	-0.0137
		(-2.69)	(-0.62)
Firm Fixed Effects	N	N	N
Year Fixed Effects	Y	Y	Y
R-squared	0.006	0.012	0.007
N	31,173	31,173	31,173

TABLE 4
NYSE Autoquote Natural Experiment

Table 4 presents the results of the NYSE Autoquote quasi-natural experiment. Panel A presents the results of the propensity model, where the dependent variable is an indicator equal to 1 if the firm is an NYSE firm, and 0 otherwise. Panel B presents the post-matching covariate balance. Panel C presents the results of the differences in differences regression, where the dependent variable is *GuideCountEPS* or *GuideCount*. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Panel A: Treatment Model

Variable	(1) *Treated*
LnMktCap	1.2294***
	(11.15)
TobinQ	-0.5224***
	(-5.19)
UE	-0.0067
	(-0.26)
LnCAR	0.8461
	(1.21)
Volatility	-34.92***
	(-3.66)
LnAnalystFollowing	-0.2654*
	(-1.77)
GuideCountEPS	0.1465**
	(2.06)
GuiderEPS	-0.0636
	(-0.29)
GuideCount	-0.1278**
	(-2.35)
Guider	-0.1315
	(-0.60)
ROA	8.216**
	(2.31)
Spread	70.99***
	(7.68)

TABLE 4, *continued*
NYSE Autoquote Natural Experiment

Panel A regression continued from prior page.

Turnover	-0.0352**
	(-2.21)
Price	0.0075
	(1.01)
Constant	-5.8134***
	(-8.84)
Industry FE	Y
R-squared	0.500
N	2,252

TABLE 4, *continued*

NYSE Autoquote Natural Experiment

Panel B: Covariate Balance

	Treated		**Control**			
Variable	Mean	Std	Mean	Std	Difference	T-statistic
PropensityScore	0.4929	0.0071	0.4885	0.0070	0.0044	(0.45)
GuideCount	0.7689	0.0343	0.8293	0.0336	-0.0604	(-1.26)
Guider	0.3995	0.0135	0.4305	0.0136	-0.0310	(-1.62)
GuideCountEPS	0.5921	0.0257	0.6405	0.0258	-0.0484	(-1.33)
GuiderEPS	0.3731	0.0133	0.4041	0.0135	-0.031	(-1.63)
LnMktCap	6.513	0.0344	6.374	0.0490	0.1390**	(2.33)
TobinQ	1.2972	0.0389	1.2009	0.0303	0.0963*	(1.95)
UE	-0.059	0.0864	-0.038	0.0635	-0.0210	(-0.19)
LnCAR	0.0102	0.0064	0.0015	0.0061	0.0087	(0.99)
Volatility	0.0293	0.0004	0.0295	0.0004	-0.0002	(0.34)
LnAnalystFollowing	1.6861	0.0238	1.5861	0.0257	0.1000***	(-2.85)
ROA	0.0081	0.001	0.009	0.0008	-0.0009	(-0.72)
Spread	0.0115	0.0002	0.0121	0.0005	-0.0006	(-0.98)
Turnover	6.9681	0.2167	6.8783	0.1939	0.0898	(0.31)
Price	21.941	0.3766	20.791	0.3866	1.1499	(0.56)

TABLE 4, *continued*
NYSE Autoquote Natural Experiment

Panel C: Regression

Variable	(1) *GuideCount*	(2) *GuideCountEPS*
Treated	-0.1560***	-0.1207***
	(-3.08)	(-3.32)
Post	0.2362***	0.0146
	(4.95)	(0.45)
*Treated * Post*	0.1177*	0.0944**
	(1.70)	(2.04)
Controls	Y	Y
Industry FE	Y	Y
R-squared	0.214	0.213
N	5,296	5,296

TABLE 5
Inverse Propensity Reweighting

This table presents the inverse-propensity reweighting with regression adjustment results. The dependent variable is *GuideCount* . The regressions compare the average treatment effect (ATE) of *AT* on *GuideCount* using a weighted sample based on the propensity to be in a given AT Quintile. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test.

AT Quintile 1

Comparison	Prediction	ATE	t-stat
2 v 1	(+)	0.538***	6.48
3 v 1	(+)	0.633***	6.73
4 v 1	(+)	1.013***	3.22
5 v 1	(+)	0.616***	2.57
N = 19,327			

AT Quintile 2

Comparison	Prediction	ATE	t-stat
1 v 2	(-)	-0.620***	-5.08
3 v 2	(+)	0.221***	3.61
4 v 2	(+)	0.313***	3.69
5 v 2	(+)	0.631***	2.84
N = 27,208			

AT Quintile 3

Comparison	Prediction	ATE	t-stat
1 v 3	(-)	-1.343***	-9.43
2 v 3	(-)	-0.484***	-5.92
4 v 3	(+)	0.144**	2.10
5 v 3	(+)	0.232*	1.79
N = 30,620			

AT Quintile 4

Comparison	Prediction	ATE	t-stat
1 v 4	(-)	-1.338***	-6.71
2 v 4	(-)	-0.631***	-4.15
3 v 4	(-)	-0.146**	-2.33
5 v 4	(+)	0.048	0.47
N = 30,524			

AT Quintile 5

Comparison	Prediction	ATE	t-stat
1 v 5	(-)	-0.695***	-4.13
2 v 5	(-)	-0.455***	-4.65
3 v 5	(-)	-0.117*	-1.75
4 v 5	(-)	0.209***	3.40
N = 25,246			

TABLE 6
Forecast Informativeness

Table 6 presents the test of AT and forecast response coefficients. The dependent variable is *CAR_Forecast* , the abnormal returns in the forecast window. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Variable	(1) *CAR_Forecast*
AT	-0.0012**
	(-2.37)
ForecastSurprise	0.0018**
	(2.29)
*AT * ForecastSurprise*	0.0004**
	(1.98)
LnMktCap	-0.0001
	(-0.16)
AnalystFollowing	-0.0000
	(-0.07)
CAR_Pre	0.0119
	(0.85)
Turnover	-0.0002
	(-1.41)
Volatility	0.0020
	(0.02)
Constant	0.0084
	(0.85)
Year Fixed Effects	Y
R-squared	0.006
N	23,543

TABLE 7
AT and Information Acquisition

Table 7 presents the results of tests on information acquisition and AT. Panel A presents the association between AT and EDGAR downloads. Panel B presents the association between AT and the magnitude of analyst forecast revisions. Panel C presents the association between stale analayst forecasts and market returns to the earnings announcement. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Panel A: EDGAR Downloads

Variable	(1) *NonRobot Downloads*	(2) *Robot Downloads*	(3) *Lead NonRobot Downloads*	(4) *Lead Robot Downloads*
AT	-1.4303***	1.5267	-0.6311**	3.4611
	(-4.12)	(0.49)	(-2.07)	(1.19)
NewFilings	0.0582	1.4508		
	(1.61)	(1.53)		
LeadNewFilings			0.0888*	1.8356
			(1.81)	(1.64)
Controls	Y	Y	Y	Y
Firm Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
R-squared	0.891	0.691	0.912	0.701
N	21,764	21,764	21,516	21,516

TABLE 7, *continued*
AT and Information Acquisition

Panel B: Analyst Revisions

Variable	(1) *AbsFrcstChange*	(2) *Lead AbsFrcstChange*
AT	-0.0003***	-0.0001**
	(-3.28)	(-2.28)
AbsUE_BegQtr	0.1694***	
	(3.96)	
LeadAbsUE_BegQtr		0.1682***
		(3.88)
Controls	Y	Y
Year Fixed Effects	Y	Y
R-squared	0.627	0.625
N	27,367	26,790

TABLE 7, *continued*
AT and Information Acquisition

Panel C: Updating Earnings Expectations

Variable	(1) *CAR_Ann*
AT	-0.0015***
	(-3.67)
UE_BegQtr	0.5289***
	(4.97)
*AT * UE_BegQtr*	0.1091***
	(4.10)
Controls	Y
Firm Fixed Effects	N
Year Fixed Effects	Y
R-squared	0.030
N	34,297

TABLE 8
Information Processing AT

Table 8 the tests the association between types of AT that are more likely to improve price informativeness and increase future disclosures. The dependent variable is *Guider*. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Variable	(1) *Guider*	(2) *Guider*
AT	0.204***	0.248***
	(4.60)	(4.85)
NewsVolume	-0.001**	-0.000**
	(-2.35)	(-2.03)
*AT * NewsVolume*	-0.000*	
	(-1.75)	
LagAT_Announcement		-0.086**
		(-2.17)
Controls	Y	Y
Firm Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
R-squared	0.023	0.023
N	14,345	13,222

TABLE 9
Illiquidity

Table 9 presents the test of AT and the cost of informed trading. The dependent variable is *Illiquidity* , the ratio of absolute stock returns to volume, measured daily and averaged over the pre-announcement period.. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

	(1)
Variable	*Illiquidity*
AT	0.244**
	(2.37)
PPUE	-0.454
	(-0.92)
LnMktCap	-0.507*
	(-1.82)
InsiderSales	0.000
	(0.00)
AnalystFollowing	0.020
	(1.33)
Dispersion	-0.126
	(-1.37)
Spread	9.098
	(0.90)
Turnover	-0.025
	(-1.46)
Volatility	44.813
	(1.47)
Skewness	-0.083*
	(-1.77)
NewsVolume	0.000*
	(1.71)
Year Fixed Effects	Y
Firm Fixed Effects	Y
R-squared	0.211
N	34,310

TABLE 10
Liquidity Making and Taking AT

Table 9 presents the results of active and passive AT on disclosure. The dependent variable is *Guider*. Panel A presents the aggregate proxies for active (*LiquidityTakingAT*) and passive (*LiquidityMakingAT*) AT. Panel B presents the individual AT proxies. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Panel A: Liquidity Making and Liquidity Taking

Variable	Prediction	(1) *Guider*	(2) *Guider*	(3) *Guider*
LiquidityMakingAT	(+)	0.1529***		0.1263***
		(4.10)		(3.37)
LiquidityTakingAT	(+)		0.1869***	0.1316**
			(3.14)	(2.20)
Controls		Y	Y	Y
Firm Fixed Effects		Y	Y	Y
Year Fixed Effects		Y	Y	Y
R-squared		0.022	0.021	0.022
N		14,345	14,345	14,345

TABLE 10, *continued*
Liquidity Making and Taking AT

Panel B: Individual Proxies

Variable	Prediction	(1) *Guider*	(2) *Guider*	(3) *Guider*	(4) *Guider*
OddLotVolumeRatio	(+)	0.2636**			
		(2.02)			
CancelToTradeRatio	(+)		0.2757***		
			(3.16)		
TradeToOrderRatio	(-)			-0.4340***	
				(-4.31)	
TradeSize	(-)				-1.0102***
					(-3.73)
Controls		Y	Y	Y	Y
Firm Fixed Effects		Y	Y	Y	Y
Year Fixed Effects		Y	Y	Y	Y
R-squared		0.020	0.020	0.022	0.022
N		14,345	14,345	14,345	14,345

TABLE 11
Forecast Specificity

Table 11 presents the results of forecast specificity on AT. The dependent variable in column (1) is *Specificity*. The dependent variable in column (2) is *AbsFrcstRange*. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Variable	(1) *Specificity*	(2) *AbsFrcstRange*
AT	-0.0151**	-0.0014**
	(-2.53)	(-2.07)
Controls	Y	Y
Firm Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
R-squared	0.293	0.563
N	15,205	13,300

TABLE 12
Disclosure Timing

Table 12 presents the results of regressing five proxies for disclosure timeliness on AT. The dependent variables are listed in the column headers. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

	(1)	(2)	(3)	(4)	(5)
Variable	*EADelay*	*FilingDelay*	*Horizon_EPS*	*Bundle_EPS*	*PostClose*
AT	-0.0930	0.4014	2.5798	-0.0030	-0.0032
	(-1.39)	(0.81)	(1.21)	(-0.72)	(-1.20)
Controls	Y	Y	Y	Y	Y
Firm Fixed Effects	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
R-squared	0.752	0.959	0.506	0.317	0.323
N	34,310	30,638	15,205	15,205	34,310

TABLE 13
Disclosure Readability

Table 13 presents the results of regressing two proxies for annual and quarterly report readability on AT. The dependent variables are listed in the column headers. See Appendix A for variable definitions. *, **, and *** indicate significance at 0.10, 0.05, and 0.01, respectively, based on a two-tailed test. Standard errors are clustered by firm.

Variable	(1) *Readability*	(2) *FileSize*
AT	0.0220	0.1013
	(1.23)	(1.31)
Controls	Y	Y
Firm Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
R-squared	0.752	0.959
N	34,310	30,638

REFERENCES

Ajinkya, Bipin B., and Michael J. Gift. 1984. "Corporate Managers' Earnings Forecasts and Symmetrical Adjustments of Market Expectations." *Journal of Accounting Research* 22 (2): 425–44. doi:10.2307/2490657.

Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5 (1): 31–56. doi:10.1016/S1386-4181(01)00024-6.

Balakrishnan, Karthik, Mary Brooke Billings, Bryan Kelly, and Alexander Ljungqvist. 2014. "Shaping Liquidity: On the Causal Effects of Voluntary Disclosure." *The Journal of Finance* 69 (5): 2237–78.

Bamber, Linda Smith, and Youngsoon Susan Cheon. 1998. "Discretionary Management Earnings Forecast Disclosures: Antecedents and Outcomes Associated with Forecast Venue and Forecast Specificity Choices." *Journal of Accounting Research* 36 (2): 167–90. doi:10.2307/2491473.

Bamber, Linda Smith, John (Xuefeng) Jiang, and Isabel Yanyan Wang. 2010. "What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure." *The Accounting Review* 85 (4): 1131–62.

Bergman, Nittai K., and Sugata Roychowdhury. 2008. "Investor Sentiment and Corporate Disclosure." *Journal of Accounting Research* 46 (5): 1057–83.

Beyer, Anne, Daniel A. Cohen, Thomas Z. Lys, and Beverly R. Walther. 2010. "The Financial Reporting Environment: Review of the Recent Literature." *Journal of Accounting and Economics* 50 (2–3): 296–343.

Billings, Mary Brooke, and Matthew C. Cedergren. 2015. "Strategic Silence, Insider Selling and Litigation Risk." *Journal of Accounting and Economics* 59 (2–3): 119–42.

Billings, Mary Brooke, Robert Jennings, and Baruch Lev. 2015. "On Guidance and Volatility." *Journal of Accounting and Economics* 60 (2–3): 161–80.

Black, Bernard S., José-Antonio Espín-Sánchez, Eric French, and Kate Litvak. 2013. "The Effect of Health Insurance on Near-Elderly Health and Mortality." SSRN Scholarly Paper ID 2103669. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2103669.

Boone, Audra L., and Joshua T. White. 2015. "The Effect of Institutional Ownership on Firm Transparency and Information Production." *Journal of Financial Economics* 117 (3): 508–33.

Bova, Francesco, Yiwei Dou, and Ole-Kristian Hope. 2015. "Employee Ownership and Firm Disclosure." *Contemporary Accounting Research* 32 (2): 639–73.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. "High-Frequency Trading and Price Discovery." *Review of Financial Studies* 27 (8): 2267–2306.

Bushman, Robert M. 1991. "Public Disclosure and the Structure of Private Information Markets." *Journal of Accounting Research* 29 (2): 261–76.

Busso, Matias, John DiNardo, and Justin McCrary. 2014. "New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators." *Review of Economics and Statistics* 96 (5): 885–97.

Carrion, Allen. 2013. "Very Fast Money: High-Frequency Trading on the NASDAQ." *Journal of Financial Markets*, High-Frequency Trading, 16 (4): 680–711. doi:10.1016/j.finmar.2013.06.005.

Chaboud, Alain P., Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. 2014. "Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market." *The Journal of Finance* 69 (5): 2045–84. doi:10.1111/jofi.12186.

Chakrabarty, Bidisha, Pamela C. Moulton, and Xu (Frank) Wang. 2016. "Attention Effects in a High-Frequency World." SSRN Scholarly Paper ID 2682673. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2682673.

Chen, Shuping, Mark L. DeFond, and Chul W. Park. 2002. "Voluntary Disclosure of Balance Sheet Information in Quarterly Earnings Announcements." *Journal of Accounting and Economics* 33 (2): 229–51.

Cheng, Qiang, and Kin Lo. 2006. "Insider Trading and Voluntary Disclosures." *Journal of Accounting Research* 44 (5): 815–48.

Cohen, Lauren, Christopher Malloy, and Lukasz Pomorski. 2012. "Decoding Inside Information." *The Journal of Finance* 67 (3): 1009–43.

Conrad, Jennifer, Sunil Wahal, and Jin Xiang. 2015. "High-Frequency Quoting, Trading, and the Efficiency of Prices." *Journal of Financial Economics* 116 (2): 271–91.

Degryse, Hans, Geoffrey Tombeur, and Gunther Wuyts. 2015. "Two Shades of Opacity: Hidden Orders versus Dark Trading." SSRN Scholarly Paper ID 2669447. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2669447.

Dye, Ronald A. 1985. "Disclosure of Nonproprietary Information." *Journal of Accounting Research* 23 (1): 123–45.

———. 1986. "Proprietary and Nonproprietary Disclosures." *The Journal of Business* 59 (2): 331–66.

———. 1998. "Investor Sophistication and Voluntary Disclosures." *Review of Accounting Studies* 3 (3): 261–87.

Egginton, Jared F., Van Ness, Bonnie F, Van Ness, and Robert A. 2014. "Quote Stuffing." SSRN Scholarly Paper ID 1958281. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=1958281.

Francis, Jennifer, Donna Philbrick, and Katherine Schipper. 1994. “Shareholder Litigation and Corporate Disclosures.” *Journal of Accounting Research* 32 (2): 137–64. doi:10.2307/2491279.

Frino, Alex, Tina Viljoen, George H. K. Wang, P. Joakim Westerholm, and Hui Zheng. 2012. “Are Algorithmic Trades Informed? - An Empirical Analysis of Algorithmic Trading Around Earnings Announcements.” SSRN Scholarly Paper ID 2132568. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2132568.

Gao, Pingyang, and Pierre Jinghong Liang. 2013. “Informational Feedback, Adverse Selection, and Optimal Disclosure Policy.” *Journal of Accounting Research* 51 (5): 1133–58.

Goldstein, Michael A., Pavitra Kumar, and Frank C. Graves. 2014. “Computerized and High-Frequency Trading.” *Financial Review* 49 (2): 177–202.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal. 2005. “The Economic Implications of Corporate Financial Reporting.” *Journal of Accounting and Economics* 40 (1–3): 3–73.

Groß-Klußmann, Axel, and Nikolaus Hautsch. 2011. “When Machines Read the News: Using Automated Text Analytics to Quantify High Frequency News-Implied Market Reactions.” *Journal of Empirical Finance* 18 (2): 321–40.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. “On the Impossibility of Informationally Efficient Markets.” *The American Economic Review* 70 (3): 393–408.

Hasbrouck, Joel, and Gideon Saar. 2013. “Low-Latency Trading.” *Journal of Financial Markets,* High-Frequency Trading, 16 (4): 646–79.

Healy, Paul M., Amy P. Hutton, and Krishna G. Palepu. 1999. “Stock Performance and Intermediation Changes Surrounding Sustained Increases in Disclosure*.” *Contemporary Accounting Research* 16 (3): 485–520.

Healy, Paul M, and Krishna G Palepu. 2001. “Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature.” *Journal of Accounting and Economics* 31 (1–3): 405–40. doi:10.1016/S0165-4101(01)00018-0.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld. 2011. “Does Algorithmic Trading Improve Liquidity?” *The Journal of Finance* 66 (1): 1–33.

Hendershott, Terrence, and Ryan Riordan. 2013. “Algorithmic Trading and the Market for Liquidity.” *Journal of Financial and Quantitative Analysis* 48 (4): 1001–1024.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.” *Econometrica* 71 (4): 1161–89.

Hirschey, Nicholas. 2016. “Do High-Frequency Traders Anticipate Buying and Selling Pressure?” SSRN Scholarly Paper ID 2238516. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2238516.

Hoffmann, Peter. 2014. “A Dynamic Limit Order Market with Fast and Slow Traders.” *Journal of Financial Economics* 113 (1): 156–69.

Huddart, Steven, Bin Ke, and Charles Shi. 2007. “Jeopardy, Non-Public Information, and Insider Trading around SEC 10-K and 10-Q Filings.” *Journal of Accounting and Economics* 43 (1): 3–36.

Imbens, G. W. 2000. “The Role of the Propensity Score in Estimating Dose-Response Functions.” *Biometrika* 87 (3): 706–10.

Kim, Irene, and Douglas J. Skinner. 2012. “Measuring Securities Litigation Risk.” *Journal of Accounting and Economics* 53 (1–2): 290–310.

King, Ronald, Grace Pownall, and Gregory Waymire. 1990. "Expectations adjustment via timely management forecasts: Review, synthesis, and suggestions for future research." *Journal of accounting Literature* 9.1: 113-144.

Kirilenko, Andrei A., Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun. 2015. “The Flash Crash: The Impact of High Frequency Trading on an Electronic Market.” SSRN Scholarly Paper ID 1686004. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=1686004.

Korajczyk, Robert A., Deborah J. Lucas, and Robert L. McDonald. 1991. “The Effect of Information Releases on the Pricing and Timing of Equity Issues.” *Review of Financial Studies* 4 (4): 685–708. doi:10.1093/rfs/4.4.685.

Korajczyk, Robert A., and Dermot Murphy. 2016. “High Frequency Market Making to Large Institutional Trades.” SSRN Scholarly Paper ID 2567016. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2567016.

Kothari, S. P., Susan Shu, and Peter D. Wysocki. 2009. “Do Managers Withhold Bad News?” *Journal of Accounting Research* 47 (1): 241–76.

Lang, Mark, and Russell Lundholm. 1993. “Cross-Sectional Determinants of Analyst Ratings of Corporate Disclosures.” *Journal of Accounting Research* 31 (2): 246–71. doi:10.2307/2491273.

Leone, Andrew J., Steve Rock, and Michael Willenborg. 2007. “Disclosure of Intended Use of Proceeds and Underpricing in Initial Public Offerings.” *Journal of Accounting Research* 45 (1): 111–53. doi:10.1111/j.1475-679X.2006.00229.x.

Leuz, Christian, and Robert E. Verrecchia. 2000. “The Economic Consequences of Increased Disclosure.” *Journal of Accounting Research* 38: 91–124. doi:10.2307/2672910.

Li, Feng. 2008. “Annual Report Readability, Current Earnings, and Earnings Persistence.” *Journal of Accounting and Economics*, Economic Consequences of Alternative Accounting Standards and Regulation, 45 (2–3): 221–47.

Li, Yinghua, and Liandong Zhang. 2015. "Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence from a Natural Experiment." *Journal of Accounting Research* 53 (1): 79–117.

Loughran, Tim, and Bill Mcdonald. 2014. "Measuring Readability in Financial Disclosures." *The Journal of Finance* 69 (4): 1643–71.

Loughran, Tim, and Bill McDonald. 2016. "The Use of EDGAR Filings by Investors." SSRN Scholarly Paper ID 2493941. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2493941.

Malinova, Katya, Andreas Park, and Ryan Riordan. 2013. "Do Retail Traders Suffer from High Frequency Traders?" SSRN Scholarly Paper ID 2183806. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2183806.

Menkveld, Albert J. 2014. "High-Frequency Traders and Market Structure." *Financial Review* 49 (2): 333–44.

Menkveld, Albert J., and Marius A. Zoican. 2015. "Need for Speed? Exchange Latency and Liquidity." SSRN Scholarly Paper ID 2442690. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2442690.

Nagar, Venky, Dhananjay Nanda, and Peter Wysocki. 2003. "Discretionary Disclosure and Stock-Based Incentives." *Journal of Accounting and Economics* 34 (1–3): 283–309.

Noe, Christopher F. 1999. "Voluntary Disclosures and Insider Transactions." *Journal of Accounting and Economics* 27 (3): 305–26.

O'Hara, Maureen. 2015. "High Frequency Market Microstructure." *Journal of Financial Economics* 116 (2): 257–70.

O'Hara, Maureen, Chen Yao, and Mao Ye. 2014. "What's Not There: Odd Lots and Market Data." *The Journal of Finance* 69 (5): 2199–2236.

Penno, Mark C. 1997. "Information Quality and Voluntary Disclosure." *The Accounting Review* 72 (2): 275–84.

Rogers, Jonathan L. 2008. "Disclosure Quality and Management Trading Incentives." *Journal of Accounting Research* 46 (5): 1265–96.

Rogers, Jonathan L., Douglas J. Skinner, and Sarah L. C. Zechman. 2016. "Run EDGAR Run: SEC Dissemination in a High-Frequency World." SSRN Scholarly Paper ID 2513350. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2513350.

Rogers, Jonathan L., and Phillip C. Stocken. 2005. "Credibility of Management Forecasts." *The Accounting Review* 80 (4): 1233–60.

Rogers, Jonathan L., and Andrew Van Buskirk. 2009. “Shareholder Litigation and Changes in Disclosure Behavior.” *Journal of Accounting and Economics*, Accounting Research on Issues of Contemporary Interest, 47 (1–2): 136–56.

Rogers, Jonathan L., and Andrew Van Buskirk. 2013. “Bundled Forecasts in Empirical Accounting Research.” *Journal of Accounting and Economics* 55 (1): 43–65.

Scholtus, Martin, Dick van Dijk, and Bart Frijns. 2014. “Speed, Algorithmic Trading, and Market Quality around Macroeconomic News Announcements.” *Journal of Banking & Finance* 38 (January): 89–105.

Schrand, Catherine M., and Robert E. Verrecchia. 2005. “Information Disclosure and Adverse Selection Explanations for IPO Underpricing.” SSRN Scholarly Paper ID 316824. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=316824.

Securities and Exchange Commission. 2010. “Securities Exchange Act Release No. 34-61358, 75 FR 3594, 3606.”

Skinner, Douglas J. 1994. “Why Firms Voluntarily Disclose Bad News.” *Journal of Accounting Research* 32 (1): 38–60.

Skinner, Douglas J. 1997. “Earnings Disclosures and Stockholder Lawsuits.” *Journal of Accounting and Economics* 23 (3): 249–82.

Sletten, Ewa. 2011. “The Effect of Stock Price on Discretionary Disclosure.” *Review of Accounting Studies* 17 (1): 96–133.

Teoh, Siew Hong, and Chuan Yang Hwang. 1991. “Nondisclosure and Adverse Disclosure as Signals of Firm Value.” *Review of Financial Studies* 4 (2): 283–313.

Verrecchia, Robert E. 1983. “Discretionary Disclosure.” *Journal of Accounting and Economics* 5 (January): 179–94.

———. 1990. “Information Quality and Discretionary Disclosure.” *Journal of Accounting and Economics* 12 (4): 365–80.

Wagenhofer, Alfred. 1990. “Voluntary Disclosure with a Strategic Opponent.” *Journal of Accounting and Economics* 12 (4): 341–63.

Weller, Brian M. 2016. “Efficient Prices at Any Cost: Does Algorithmic Trading Deter Information Acquisition?” SSRN Scholarly Paper ID 2662254. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2662254.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Yang, Liyan, and Haoxiang Zhu. 2016. “Back-Running: Seeking and Hiding Fundamental Information in Order Flows.” SSRN Scholarly Paper ID 2583915. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2583915.

Zhang, Guochang. 2001. “Private Information Production, Public Disclosure, and the Cost of Capital: Theory and Implications.” *Contemporary Accounting Research* 18 (2): 363–84.

Zhang, S. Sarah. 2012. “Need for Speed: An Empirical Analysis of Hard and Soft Information in a High Frequency World.” SSRN Scholarly Paper ID 1985951. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=1985951.

APPENDIX A

Variable Definitions

AT Proxies	**Definition**
AT	The first principal component from the *OddLotVolumeRatio, TradeToOrderRatio, CancellationToTradeRatio,* and *TradeSize.*
AT_Change	The first difference of *AT* from the prior quarter.
CancellationToTradeRatio	Number of orders cancelled divided by the number of trades executed. Measured daily and averaged over the preannouncement period.
LiquidityMakingAT	The first principal component from the *TradeToOrderRatio* and *CancelToTradeRatio,* adjusted so it is positively correlated with AT.
LiquidityTakingAT	The first principal component from the *OddLotVolumeRatio* and *TradeSize* , adjusted so it is positively correlated with AT.
OddLotVolumeRatio	Volume executed in odd-lot trades divided by the total volume traded. Measured daily and averaged over the preannouncement period.
TradeSize	Total volume executed divided by the number of trades placed, multiplied by -1. Measured daily and averaged over the preannouncement period.
TradeToOrderRatio	Total volume executed divided by total volume in orders placed, multiplied by -1. Measured daily and averaged over the preannouncement period.

Disclosure Variables	
AbsFrcstRange	The absolute value of the range of the EPS forecast with the longest horizon in the disclosure period, scaled by price at the beginning of the preannouncement period.
Bundle_EPS	An indicator equal to 1 if the firm's longest horizon EPS forecast is bundled with the earnings announcement.
EADelay	The number of days between the period end date and the earnings announcement date.
FileSize	The file size of the firm's annual or quarterly report, scaled by 1,000,000.
FilingDelay	The number of days between the earnings announcement date and the associated 10-K or 10-Q filing.
GuideCount	A count of the pieces of guidance issued during the disclosure period.
GuideCount_Ann	The count of guidance issued during the [-1,+1] earnings announcement window.
GuideCount_Pre	The count of guidance issued during the preannouncement period.
GuideCount_Pre2	The count of guidance issued during the [-126,+5] window around the prior earnings announcement.

GuideDays	The number of distinct days on which the firm issued guidance during the disclosure period.
Guider	An indicator equal to 1 if the firm issues any guidance in the disclosure period, 0 otherwise.
Horizon_EPS	The number of days between the forecast date and the date forecasted. Measured for the EPS forecast with the longest horizon in the disclosure period.
PostClose	An indicator equal to 1 if the firm announces earnings outside of normal trading hours, otherwise 0.
Readability	The first principal component of the eight readability proxies available in the WRDS SEC Analytics Suite from the firm's annual or quarterly report.
Specificity	Equals 3 if the forecast is a point estimate, 2 if the forecast is a closed range, and 1 if the forecast is an open ended range. Measured for the EPS forecast with the longest horizon in the disclosure period.

Other Variables

AbsFrcstChange	The absolute value of the difference between the consensus forecast measured at the beginning of the preannouncement period and the consensus forecast measured at the end of the preannouncement period, scaled by price.
AbsUE_BegQtr	The absolute value of *UE_BegQtr*.
Accruals	Earnings less operating cash flows, before extraordinary items and discontinued operations, scaled by earnings.
AnalystFollowing	Number of analysts with a forecast in the consensus forecast calculation.
Annual	An indicator equal to 1 if the current quarter is the last quarter of the fiscal year, 0 otherwise.
CAR_Ann	Cumulative abnormal returns during the three day earnings announcement window [-1.+1], calculated using the Fama-French three factor model.
CAR_Forecast	Cumulative abnormal returns during the two day forecast announcement window [0.+1], calculated using the Fama-French three factor model.
CAR_Pre	Cumulative abnormal returns for the firm during the preannouncement period, calculated from the Fama-French three factor model.
Dispersion	The standard deviation of the forecasts used in the consensus forecast calculation.
Dividends	An indicator variable equal to 1 if the firm paid any dividends during the fiscal year, 0 otherwise.
ForecastSurprise	The EPS forecast issued by management less the analyst consensus mean, scaled by the analyst consensus.

FrcstChangeStd	The standard deviation of *AbsFrcstChange* by firm over the sample period.
Illiquidity	The absolute value of returns divided by total volume traded, measured by day and averaged over the pre-announcement period.
InsiderSales	The net insider sales and purchases by executives and board members during the preannouncement period (positive number indicates net sales).
LeadAbsFrcstChange	The following quarter's value of *AbsFrcstChange* .
LeadAbsUE_BegQtr	The following quarter's value of *AbsUE_BegQtr*.
LeadNewFilings	The following quarter's value of *NewFilings*.
LeadNonRobotDownloads	The number of non-robot downloads of SEC filings for a given firm measured from the day before the current earnings announcement to two days prior to the next earnings announcement.
LeadRobotDownloads	The number of robot downloads of SEC filings for a given firm measured from the day before the current earnings announcement to two days prior to the next earnings announcement.
Leverage	The leverage of the firm, calculated as short term debt plus long term debt divided by total assets.
LnMktCap	Natural log of the market capitalization of the firm at the beginning of the preannouncement period.
Loss	An indicator equal to 1 if earnings is negative, 0 otherwise.
NegUE	An indicator equal to 1 if UE < -0.0001, 0 otherwise.
NewFilings	The number of new SEC filings made by a firm in the window from one day before the prior earnings announcement to two days before the current earnings announcement.
NewsVolume	The aggregate event volume (AEV) variable from RavenPack that calculates a score for the level of news in the prior 91 days using the quantity and novelty of news articles written about the firm. I keep the AEV closest but prior to the upcoming earnings announcement.
NonRobotDownloads	The number of non-robot downloads of SEC filings for a given firm in the preannouncement period.
PosUE	An indicator equal to 1 if UE > +0.0001, 0 otherwise.
PPUE	The percentage of the four prior quarters with with positive unexpected earnings (*PosUE* = 1).
RobotDownloads	The number of robot downloads of SEC filings for a given firm in the preannouncement period.
SalesGrowth	The percentage change in total revenue from the prior quarter.
Skewness	Skewness of returns during the window [-126,+5] around the lagged earnings announcement date.

Spread	The average bid-ask spread, scaled by price, during the window [-126,+5] around the lagged earnings announcement date.
Turnover	Average turnover (volume divided by shares outstanding) per day during the window [-126,+5] around the lagged earnings announcement date.
UE	Actual earnings from IBES less the consensus analyst forecast, scaled by stock price at the beginning of the preannouncement period.
UE_BegQtr	Unexpected earnings calculated the same as *UE,* but instead measuring the consensus forecast five days following the prior earnings
Volatility	Standard deviation of returns during the window [-126,+5] around the lagged earnings announcement date.

NYSE Autoquote Experiment Variables

GuideCountEPS	The same as *GuideCount* , but only for EPS forecasts.
GuiderEPS	The same as *Guider* , but only for EPS forecasts.
Price	The average price of the firm in the preannouncement period.
ROA	Earnings divided by total assets.
TobinQ	The market capitalization of the firm divided by total assets.
Treated	An indicator equal to 1 if the firm is listed on the NYSE, otherwise 0.
Volatility	The standard deviation of the natural log of cumulative abnormal returns over the pre-treatment period.