DOES PUBLIC INFORMATION ACQUISITON LEVEL THE PLAYING FIELD OR WIDEN THE GAP? AN ANALYSIS OF LOCAL AND NON-LOCAL INVESTORS

Travis A. Dyer

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Accounting Department in the Kenan-Flagler Business School.

Chapel Hill
2018

Approved By:

Mark Lang

Alon Brav

Robert Bushman

Eva Labro

Wayne Landsman

ABSTRACT

Travis A. Dyer: Does Public Information Acquisition Level the Playing Field
or Widen the Gap? An Analysis of Local and Non-Local Investors
(Under the direction of Mark Lang)

While prior studies have speculated that local investors have an information advantage as a result of private information, I examine whether local investors generate an information advantage through their use of *public* information. Using novel data on professional investors' requests for financial information from the SEC, I document that investors acquire more financial information for their local investments, and that local investors are more timely at acquiring newly released disclosures. Consistent with public information being more valuable to local investors, I find that investors make more profitable trading decisions in local stocks when they also acquire public financial information. Lastly, I find that investors' preference to over-weight local stocks in their investment portfolio ("Local Bias") is increasing in the amount of public information acquired. These results suggest that the use of public information may widen the information gap between local and non-local investors rather than level the playing field.

ACKNOWLEDGEMENTS

I wish to thank members of my dissertation committee, Alon Brav, Robert Bushman, Eva Labro, and Wayne Landsman for their helpful suggestions, and especially Mark Lang (chairman) for his mentorship and insights. I have also benefited from the helpful comments of Andrew Belnap, Michael Drake, Dimas Pena Romera, Lorien Stice-Lawrence, Donny Zhao, conference participants at the 2017 BYU Accounting Research Symposium, and workshop participants at The University of North Carolina. All mistakes are my own.

TABLE OF CONTENTS

LIST OF TABLES…….vii

LIST OF FIGURES.....viii

LIST OF ABBREVIATIONS…….ix

CHAPTER 1: INTRODUCTION……1

CHAPTER 2: DEVELOPMENT OF HYPOTHESES……..10

CHAPTER 3: DATA AND EMPIRICAL DESIGN…….14

Sample Construction and Characteristics……..14

Research Design and Variable Construction…….16

CHAPTER 4: RESULTS AND ROBUSTNESS……...22

Tests of H1: Proximity and Public Information Acquisition…….22

Main Analysis……22

Exogenous Shock to Investor Proximity……24

Tests of H2: Value of Information Acquisition for Local and Non-local Investors.....26

Enhanced Timeliness…….26

Improved Decision Making……...27

Tests of H3: "Local Bias" and Information Acquisition……...28

Robustness of Empirical Results……...28

Investment Level Control……..28

Familiarity Bias…….29

Controlling for Relative Wealth Hedging…….30

Information Asymmetry……30

Non-Professional Investors……31

Alternative Measure of Proximity…..……32

CHAPTER 5: CONCLUSION.…….33

APPENDIX 1: VARIABLE DEFINITIONS AND MEASUREMENT.……..36

REFERENCES…….…….38

FIGURES...…….…….……...….....41

TABLES…….…….……...….....42

LIST OF TABLES

Table 1 – Professional Investor Descriptive Statistics……...42

Table 2 – Descriptive Statistics of *Primary Sample*……..43

Table 3 – Relation between Public Information Acquisition and Investors' Proximity to Investment……44

Table 4 – Relation between Public Information Acquisition and Investors' Proximity to Investment by Information Type Requested……45

Table 5 – Cross-Sectional Variation in the Relation between Public Information Acquisition and Investors' Proximity to Investment: Relationship Age……...46

Table 6 – Relation between Public Information Acquisition and Investors' Proximity to Investment Using Changes in Firm Headquarters……47

Table 7 – Relation between the Delay in Public Information Retrieval and Investors' Proximity to Investment……48

Table 8 – Descriptive Statistics for the *13(f) Sample*……....49

Table 9 – Public Information Acquisition and Informed Trading……50

Table 10 – Relation between Local Bias and Public Information Acquisition……..51

Table 11 – Main Findings Controlling for the Level of Investment……..52

Table 12 – Firm-Level Information Asymmetry and Local Research……...53

Table 13 – Relation between Public Information Acquisition and Investors' Proximity to Investment Using Alternative Measure of Proximity……54

LIST OF FIGURES

Figure 1A – United States Zip Codes with Professional Investors using EDGAR……...41

Figure 1B – United States Zip Codes with Firm Headquarters…….41

LIST OF ABBREVIATIONS

EDGAR	Electronic Data Gathering, Analysis, and Retrieval system
IRS	Internal Revenue Service
PIA	Public Information Acquisition
SEC	Securities and Exchange Commission

CHAPTER 1: INTRODUCTION

Many claim that investors have a private information advantage in local stocks, which causes them to tilt their allocation of invested wealth towards geographically proximate investments (French and Poterba, 1991; Coval and Moskowitz, 1999; Coval and Moskowitz, 2001; Cooper et al. 2012). But what is the role of *public* information for local and non-local investors? On one hand, public information might substitute for private information advantages and therefore be used as an equalizer to reduce the private information advantages of local investors. On the other hand, if public and private information act as complements, then local investors might use public information to widen their information advantage. In this study, I examine public information acquired by professional investors to see (1) if investors acquire more public information for local or non-local investments, (2) if public information is more valuable to local or non-local investors, and (3) if public information acquisition mitigates or moderates investors' preference to over-weight local stocks in their investment portfolio (i.e., "Local Bias").[1]

In this study, I construct and utilize a direct measure of professional investor information acquisition. Until recently, research on investor information retrieval has been quite challenging to study directly (Drake et al, 2015; Drake et al, 2016; Lee et al. 2015; Loughran and McDonald, 2015). With the release of EDGAR log files by the Security and Exchange Commission (SEC),

[1] "Local Bias" is the term used in academic literature to describe the well-documented phenomenon that investors tilt their allocation of invested wealth towards geographically proximate investments in the United States (Coval and Moskowitz, 1999). The reference to bias in the term does not imply that the phenomenon is a result of a behavioral bias.

the barriers to this stream of literature have been removed. EDGAR is a large repository of all public financial information provided to and maintained by the SEC. The EDGAR log files contain each request received by the EDGAR system, documenting a partially masked IP address performing the request, the timing of request, and the filing requested.

Using ARIN WhoWas reports, I am able to unmask the identity of IP addresses that download documents from EDGAR and directly link specific professional investor institutions to their information requests on EDGAR. This process not only allows me to know where the investor is located, but also allows me to link in other investor-level data, such as investors' public 13(f) holdings. This process results in linking approximately 9,810 professional investors to EDGAR download data. The most research-active professional investor companies in my sample include companies such as JP Morgan and Citi Group. The most frequent locations of professional investors include New York City and Chicago.

The richness of the underlying data provides this study the opportunity to implement a unique and robust research design. In its simplest form, the data is a record of an investor's request for firm information. The same investor can request information for multiple firms, and the same firm can have multiple investors requesting information. Because of this feature, I can include an investor-time fixed effect, which controls for the average research activity by an investor in a given period of time. Additionally, I can include a firm-time fixed effect, which controls for the average research for a firm in a given period of time. The combination of these effects controls for attributes such as an investor's location, skill, risk appetite, and resources, as well as a firm's size, visibility, risk, and performance in that period of time. The remaining variation in research behavior is abnormal for both the investor and firm in that time period. As

such, the scope of alternative explanations for observing a proximity effect on information retrieval drastically narrows.

Using this research design and sample of professional investors, I find evidence that investors tend to download more information from EDGAR for proximate investments.[2] For example, each calendar quarter investors access 20.16% more kilobytes of information if they are within 100 kilometers of the investment.[3] Additionally, this positive relation is quite pervasive. It exists for multiple proxies of information acquisition. It persists across all major form types filed with the SEC, including both annual and quarterly reports, as well as material disclosures (i.e., 8-Ks). The local effect also persists throughout the entire investment relationship. I also find that research by local investors tends to be particularly concentrated around informational events. This evidence suggests that investors obtain more public information for local investments and is consistent with investors choosing to use public information to widen their information advantage, rather than to mitigate their information disadvantages.

One concern with the prior analyses is that an investor's choice to acquire information about a firm is endogenous to their choice to follow the firm. To address this concern, I perform an analysis using plausibly exogenous variation in investor proximity. To obtain this variation I use a sample of firms who chose to relocate their firm headquarters. Prior work suggests that firm headquarter locations are determined based on strategic business decisions related to lower taxes, lower employee wages, and better airport facilities (Strauss-Kahn and Vives, 2009). I

[2] A candidate correlated omitted variable is the level of investment in a stock. Investors might research more for stocks that make up a greater proportion of their portfolio. This is important since research confirms investors hold larger portfolio positions in local stocks (i.e., "Local Bias"). In robustness tests, I control for the level of investment and find similar results.

[3] I follow Coval and Moskowitz (2001) which uses 100 kilometers as an indicator of being proximate to an investment. As discussed in later sections, inferences are robust to alternative cutoffs and measurements.

believe that it is unlikely that the decision of where to relocate a firm's headquarters is made by executives based on the preferences of investors for proximity.[4] In this setting, if investors who exogenously become local request more information about the firm *after* the headquarter change, then proximity, rather than firm selection, is the dominant factor affecting their information acquisition decisions. With this sample of firms, I find evidence consistent with proximity to investments leading to increased information acquisition by professional investors.

I next explore the value of the observed asymmetric public information acquisition. Prior work surrounding local investors' trading profitability more generally is mixed with some finding that investors are more profitable in local investments and others the opposite (Coval and Moskowitz, 2001; Ivkovic and Weisbenner, 2005; Massa and Simonov, 2006; Seasholes, 2010; Bernile et al. 2016). One could argue that asymmetric information might exist, but that it does not provide actionable value to investors' portfolios. As such, I measure the value of public information for local investors in two ways: (1) enhanced speed of acquisition, and (2) enhanced portfolio decision-making. First, obtaining public information in advance of other investors is often valuable. For example, investors achieve trading profits through the advance access to consumer reports (Mullins et al., 2013). If local investors obtain public information more quickly than their non-local counterparts, then they have the opportunity to make use of the information more rapidly than the average investor. This speed advantage might even discourage non-local investors from performing as much stock research in the first place. I provide evidence that investors tend to acquire public information more rapidly for local investments. Looking at the two days following the release of SEC filings, I find that investors arrive at the public filing of their local investments approximately 30 minutes sooner than that of their non-local investments.

[4] This assumption is likely valid since there are multiple investors spread across the country, each possibly vying for increased proximity to the firm.

Looking at only the 60 minutes following the release, I also find that local investors still have a time advantage. Since investors have autonomy over when they access the new disclosures, these tests provide investors' revealed preference for timeliness. As such, the findings suggest investors value timeliness of information retrieval more for local stocks.

Second, a more direct measure of value is the potential benefit of public information for improving portfolio allocation decisions. I explore this possibility by linking professional investors' research on EDGAR with their 13(f) portfolio holdings data. With this data, I first confirm that more research about a stock improves investors' portfolio choices. I find that share purchases are more positively related to future returns when institutional investors perform research about a stock. Specifically, future returns tend to be 27 basis points higher (lower) when share purchases (sales) are coupled with public information research. More importantly, I find that the benefit of stock research for portfolio decisions is approximately 54 basis points larger for local stocks. This evidence suggests that public information is, on average, more valuable to local investors.

I conclude the paper by looking at the relation between "Local Bias" and public information acquisition. A large stream of literature suggests that asymmetric private information between local and non-local investors is a driving force behind investors' preference to tilt their portfolios toward local stocks, a phenomenon which prior research has dubbed "Local Bias." If public information is a contributing force to investors' "Local Bias," then an investor's "Local Bias" should vary with the amount of public information acquired about the stock. Alternatively, if "Local Bias" is the result of a non-information-based force, then "Local Bias" should be unrelated to an investor's information acquisition. Using the sample of 13(f) institutional investors, I find that an investor's "Local Bias" is increasing in the amount of public information

acquired about the stock. This suggests that a contributing force to investors' "Local Bias" is the amount of public information acquired and is consistent with information asymmetries motivating investors to allocate more wealth toward local investments.

This paper is motivated on four fronts. First, one of the regulatory objectives of the Securities and Exchange Commission (SEC) is to "level the playing field" among investors through public financial disclosure (SEC, 2000). Put into context, the SEC's objective is to equalize the information sets *available* to investors. According to the SEC, the primary purpose for providing public filings on EDGAR is to "increase the efficiency and fairness of the securities market" (SEC, 2010). This paper explores which investors benefit most from public disclosures and whether investors choose to *acquire* public information in a manner consistent with either equalizing or diverging information sets across investors.

Second, there is a large disagreement among academics about whether local investors have asymmetric information. Some contend that local investors have valuable private information (Ivkovic and Weisbenner, 2005; Massa and Simonov, 2006), while others contend no such information seems to exist (Seasholes, 2010) or largely disappeared at the turn of the century (Bernile et al., 2016). Additionally, other streams of research attribute "Local Bias" to behavioral rather than information-based explanations (Coopers et al., 2012). I speak to this ongoing debate by observing the public information sets of investors and examining whether information asymmetries exist.

Third, Cooper et al. (2012) suggests that researchers should attempt to identify the source of local investors' information advantage, should it exist. Recent theoretical work by Van Nieuwerburgh and Veldkamp (2009) suggests that *public* information may be one source of the asymmetry. Directly pertinent to this study, they predict that, under a reasonable set of

assumptions, investors will acquire more public information for local stocks if they are endowed with a small local information advantage. My paper explores this possibility by investigating whether investors acquire more public information for local or non-local stocks.

Fourth, "Local Bias" is an economically large phenomenon. Prior work on "Local Bias" finds that the typical US household allocated 30% of their portfolio to local stocks even though those same stocks make up only 12% of the total US market on average (Seasholes and Zhu, 2010). This 18% gap in US investor diversification equates to approximately $1.12 trillion or 6.4% of GDP in the year 2014.[5] Other research finds that "Local Bias" also exists for mutual fund investors (Coval and Moskowitz, 1999; Bernile et al. 2016). "Home Bias," the international equivalent to "Local Bias," is also quite large. According to Investment Company Institute (2015), only 25% of US equity mutual fund assets were in equity funds focused on foreign markets in 2014, with foreign markets encompassing 55% of the world market (Ferri et al., 2013). Therefore, understanding the extent to which public information acquisition mitigates or moderates investors' "Local Bias" is economically important.

This paper contributes to the budding literature on information retrieval. Prior research in this area studies the determinants and consequences of information retrieval at the firm level (Drake et al., 2015; Drake et al., 2016), the filing level (Lee et al., 2015; Loughran and McDonald, 2015), and the geographic level (Mondria et al., 2010; Drake et al., 2017). My study builds upon this literature by disaggregating information retrieval to the investor level. To my knowledge, this study provides the first direct test of how proximity between investors and firms relates to public information retrieval.

[5] US registered investment companies hold $15.9 trillion total mutual fund assets, with 52% in equity investments and 75% of those holdings in US equities (Investment Company Institute, 2015). Thus 1.12 = (15.9*.52*.75*.18)

More generally, this paper speaks to a broad literature on portfolio choice and profitability. Prior academic research provides convincing evidence that investors tilt their portfolio toward geographically proximate investments (Cooper et al., 2012) and suggests this might be related to asymmetric private information advantage. This study provides input-based evidence that local and non-local investors create asymmetric information sets using public information. Furthermore, this study builds on prior research that documents increased profitability of investors in local stocks (Coval and Moskowitz, 2001; Ivkovic and Weisbenner, 2005; Giannini et al., 2015; Bernile et al., 2015) and provides evidence consistent with public information being one channel through which this advantage occurs.

This paper also contributes new insights into how investors gain their local information advantage. Prior research using Twitter and stock message boards suggests that local investors have more information about local investments. Giannini et al. (2015) finds that the content in Twitter posts are more predictive of future stock performance when written by local users. Huang et al. (2012) finds that message board users in China post more about local stocks. This study complements these findings and suggests that one channel through which local investors increase their information advantage is the acquisition of public information.

While I have attempted to mitigate potential concerns, the results are subject to caveats. First, causal inferences need to be appropriately caveated. Even though the research design choices and other robustness tests mitigate several omitted-variable biases and reverse causality concerns, it is possible that alternative explanations persist. Second, while I perform several data screens and robustness tests to increase confidence that my analyses capture actual investors, there is still the possibility that my sample continues to include other users of financial

statements. Nonetheless, this analysis provides novel information on the impetus of public information acquisition.

In the next section, I discuss relevant literature and present my primary hypotheses. In Section 3, I discuss the data and methodology. I provide both empirical results and robustness analysis in Section 4. Section 5 concludes.

CHAPTER 2: DEVELOPMENT OF HYPOTHESES

Prior academic literature often speculates that local investors have a private information advantage over non-local investors. However, the role of *public* information for local and non-local investors has received little attention. The relation between public and private information in this setting is not obvious. One possibility is that local private information may act as a substitute for publicly available information. Accordingly, the incentive to acquire public information on local stocks would be mitigated by local private information. The concept of public and private information acting as substitutes is quite common. An example of this is found in a recent study of IRS public information acquisition. Bozanic et al. (2017) finds that the IRS's incentive to acquire public information from EDGAR decreased after they obtained access to private tax information. Furthermore, the SEC catch phrase "level playing field" asserts a substitutive nature between public and private information.

Studies on "Local Bias" in particular have made the argument of substitutes. Lewis (1999) highlights a main counterargument to observing information asymmetries between local and non-local investors by writing, "Greater uncertainty about foreign returns may induce the investor to pay more attention to the data." This reasoning is built on the premise that public and private information are substitutes. In recent theoretical work, Van Nieuwerburgh and Veldkamp (2009) investigate investors' decision to acquire public information. They model investors endowed with a small initial local advantage and find that, given equal access to local and non-local information (i.e., public information), a constraint on information acquisition, and a fixed portfolio, investors acquire more information about non-local investments. Since investors

cannot change their portfolio return, they prefer to acquire information that reduces the most portfolio risk (i.e., information for non-local investments).

The second possibility is that public and private information act as complements. Accordingly, the incentive to acquire public information about local stocks would be greater with local private information. Mosaic theory, for example, suggests that the combination of non-material private information with public information creates an information mosaic that is more valuable than the sum of the individual parts. Prior research has documented that under certain circumstances, public information may act complementary to private information. Research studying public disclosure of earnings has observed increases in intra-day spreads, consistent with a complementarity between public and private information (Lee et al., 1993). Van Nieuwerburgh and Veldkamp (2009) predicts that when investors have the joint decision to acquire public information as well as make portfolio investment choices they choose to acquire more information about local investments. In this setting, prices reveal the average investor's knowledge. Learning about non-local assets sacrifices excess returns, since both local and non-local investors have knowledge of future payouts. If investors instead learn about local assets, prices will only reflect as much information as the average investor knows. As the authors put it, "Investors profit more from knowing information others do not know."

Whether public information is a substitute or complement to local private information is an empirical question. Given the preceding discussion, I formulate the following hypothesis in the null form:

> ***H1:*** *The amount of public information acquired by an investor is unrelated to their proximity to investment.*

I next investigate the value of public information to local and non-local investors. Prior research surrounding local investors' trading profitability has found that local investors make either more profitable trading decisions (Coval and Moskowitz, 2001; Ivkovic and Weisbenner, 2005; Massa and Simonov, 2006), less profitable trading decisions (Seasholes, 2010) or that the profitability has decreased at the turn of the century (Bernile et al. 2016). Based on the conflicting results, one question is whether the asymmetric public information provides actionable value to an investor's portfolio. If public information is a complement [substitute] to local private information, then public information for local [non-local] investments should be more valuable to investors. However, if public information is stale or irrelevant for portfolio holdings purposes, then there may be no value to asymmetric information acquisition. Stated in the null form:

H2: *The value of public information is unrelated to an investor's proximity to the investment.*

I lastly explore the relation between "Local Bias" in invested wealth and public information acquisition. Prior academic literature finds that both mutual fund managers and individual investors exhibit "Local Bias" in equity portfolios (Coval and Moskowitz, 1999; Coval and Moskowitz, 2001; Ivkovic and Weisbenner, 2005; Massa and Simonov, 2006; Seasholes and Zhu, 2010). However, there is much debate on why this behavior exists (Cooper et al., 2012). A major theory on the topic is asymmetric private information advantages (Coval and Moskowitz, 1999; Coval and Moskowitz, 2001), but other explanations include relative wealth hedging (Massa and Simonov, 2006), transaction costs (Choe et al., 2005), familiarity (Huberman, 2001), loyalty (Cohen, 2009), patriotism (Morse and Shive, 2011), and overconfidence (Graham et al., 2009). The crucial ingredient separating asymmetric information

advantages from the other theories is the presence of private information held only by local investors.

If public information is used more by non-local investors (i.e., substitutes for private information), then it should mitigate "Local Bias" under the asymmetric information advantage theory. Alternatively, if public information is used more by local investors (i.e., complements private information), then it should further increase "Local Bias." The later explanation is consistent with Van Nieuwerburgh and Veldkamp (2009), which predicts that the use of public information increases "Local Bias." In contrast, if "Local Bias" is driven a non-information based explanation, then acquired public information should be unrelated to investors' "Local Bias." Stated in the null form:

H3: *Local Bias is unrelated to public information acquisition.*

CHAPTER 3: DATA AND EMPIRICAL DESIGN

3.1 Sample Construction and Characteristics

I use EDGAR log files to capture information acquisition by investors. Critical to my study, EDGAR log files retain the masked IP address, the filing acquired, and the time of each information request. Similar to prior studies which use EDGAR log files, I exclude (1) self-identified web crawlers, (2) IP addresses that make more than 1,000 requests in a day, (3) IP addresses that request filings for more than 100 firms in a day, (4) server codes of 300 or greater (e.g., URL not found), and (5) information requests of index files and images (Drake et al., 2015; Lee et al., 2015; Loughran and McDonald, 2015).

I next match masked IP addresses requesting EDGAR filings with IP address ownership data from ARIN WhoWas reports.[6] WhoWas reports are similar to the more common WhoIs reports, but instead provide historical ownership information for IP addresses rather than just the current ownership information. If the entire IP address block is owned by an entity in a given period, I assign that entity's name to the IP address in question.[7] I identify investors from these names by manually categorizing the 4,000 most used names in my sample. For the remaining names, I use keywords associated with investment advisors identified in the initial 4,000 names.[8]

[6] Given the limitations on ARIN WhoWas requests, I limit requests to IP address blocks with more than 10 observations in my sample.

[7] Since the last octet of IP addresses provided in the SEC log files are scrambled (e.g., 12.25.278.ajc), I refer to an "IP address block" as the range of possible IP addresses (e.g., 12.25.278.0 through 12.25.278.255).

[8] These key words include variations of the words: Asset Management, Bank, Broker, Capital, Causality, Commerce, Credit, Equity, Fiduciary, Fund, Leasing, Ledger, Lendor, Insurance, Investment, Investor, Merchant, Mortgage, Mutual, Pension, Profit, REIT, Saving, Securities, Trade, Trust, Underwriter, and Wall Street.

For the entirety of my analyses, I use the resulting sample of IP addresses categorized as investors and refer to these investors as professional investors.[9] The sample of EDGAR data employed in this study ranges from the first quarter of 2003 to the first quarter of 2015. Table 1 Panel A reports the top 10 most research active professional investor companies. As to be expected, this list includes companies such as JP Morgan, Morgan Stanley, and Wells Fargo. Table 1 Panel C reports the time-series variation in the number of professional investors identified using EDGAR. In general, the number of professional investors using EDGAR has remained relatively constant over the sample period.

I obtain location data from two sources. Investor location data is obtained from IP2Location, which links zip codes to specific IP addresses.[10,11] To obtain the zip code location of the firm whose information is being acquired, I use the header information from SEC 10-K and 10-Q filings. Specifically, I used the zip code used for the business address. If zip code data is missing or changes over the sample period, I exclude these investors and firms from the main sample.[12] As such, for the main analyses, proximity between firms and investors is time invariant. Following prior literature (Coval and Moskowitz, 1999), I exclude zip codes from Alaska, Hawaii, and all US territories.

Table 1 Panel B reports the top 10 most frequent cities for professional investor firms. I find that New York, Atlanta, and Chicago are very common cities for professional investors. This list generally aligns with priors on cities with strong financial centers. Figure 1A illustrates

[9] In later robustness tests, I used non-professional investor IP addresses in EDGAR and reach similar conclusions.

[10] I use IP2LOCATION-LITE-DB9 available at http://lite.ip2location.com.

[11] Since EDGAR masks the last octet of IP address in the publicly available log files, there are instances where it cannot be determined which zip code to assign a given IP address. I remove these IP addresses from my sample.

[12] As described later in further detail, I used some firms that change headquarter locations during the sample period for a robustness test.

the geographic distribution of professional investor EDGAR users across the United States. From this figure, we observe that the use of EDGAR seems to be spread across the United States, suggesting that EDGAR use has not been confined to just a few large cities. This is reassuring, as variation in professional investor location is vital to testing the relation between information acquisition and proximity to investment. Figure 1B illustrates the geographic distribution of company headquarters in my sample across the United States. Both investor location and company headquarters location tend to be found in or near large cities, but are relatively spread across the United States.

3.2 Research Design and Variable Construction

A major strength of this study is the research design implemented throughout each of the analyses, which capitalizes on the available counterfactual or benchmark groups using a variety of fixed effects. Since investors research multiple investments at any given point in time, an ideal benchmark for an investor's research behavior for a local firm is the same investor's research behavior for a distant firm during the same time period. Using a within investor-time analysis strengthens identification by removing possible correlated omitted variable biases relating to fixed investor attributes in that time-period (i.e., investor location, investor type, investor risk appetite, investor resources, investor wealth, etc.). In tandem with an investor-time analysis, I also incorporate a firm-time benchmark, which controls for fixed firm characteristics in that time period (i.e., firm size, firm risk, firm visibility, firm performance, the amount of information available for request, etc.). The combination of within investor-time and within firm-time analyses increases internal validity by removing a host of alternative explanations in each of the analyses. What remains after incorporating these two fixed effect structures is variation associated with the pair itself, investor and firm.

Following the approach recommended in Gormley and Matsa (2014), each of the empirical designs uses a high-dimensional fixed effects structure.[13] To test whether investors consume more or less public information for local investments (*H1)* I use the following research design:

$$PIA_{i,j,t} = \alpha_1 \cdot Local_{i,j} + \alpha_{i,t} + \alpha_{j,t} + \varepsilon_{i,j,t} \quad (1)$$

PIA represents the amount of public information acquisition by investor *i* about firm *j* in period *t* and *Local* is an indicator if the investor is local to the investment. Coefficients $\alpha_{i,t}$ and $\alpha_{j,t}$ represent the investor-time and firm-time fixed effects, respectively. *H1* is rejected in favor of the alternative hypothesis if α_1 is significantly different from zero. If α_1 is also positive [negative] then the results would be consistent with public and private information acting as complements [substitutes].

I measure the amount of public information acquired by investors (*PIA*) using four measures. The first is the total kilobyte size of files requested for the first time (*SumSize*). This is intended to capture the amount of new information an investor is acquiring during the quarter. The logic behind this measure is that the amount of information in a filing is likely related to the number of kilobytes in the filing. The second and third measures are the number of files requested for the first time (*InitialDownloads*) and the number of files requested for a second time or more (*Redownloads*). Each of these variables are measured over calendar quarters and are logged to reduce the influence of outliers.[14,15] The fourth measure is the principal component

[13] Gormley and Matsa (2014) recommends fixed effects over adding means of dependent variables as controls. They further recommend certain Stata functions when faced with high-dimensional fixed effects. This paper uses reghdfe. Additional information is available on their website: http://finance.wharton.upenn.edu/~tgormley/papers/fe.html.

[14] The volume of EDGAR requests is such that evaluating inferences on a quarterly basis, instead of a weekly or daily basis, makes estimation more computationally feasible.

[15] Results are robust to using the unlogged variables.

of the previous three measures (*Factor*).[16] For all of the above information acquisition variables, interim periods where professional investors do not acquire any information about a company they follow are categorized as zero.[17] *Local* is an indicator variable set to one if the geodesic distance between the investor zip code and the firm headquarters zip code (i.e., as the crow flies) is less than 100 kilometers (Coval and Moskowitz, 2001).[18] All continuous variables are winsorized at the 1 and 99 percent level by quarter. Standard errors are clustered by investor and firm.

I test the value of public information for local and non-local investors (*H2)* using two approaches. The first approach focuses on the timing of information acquisition. An enhanced speed in acquiring information affords investors the ability to make more timely trading decisions on new information. In other words, speed is valuable to investors. If investors tend to acquire filings more rapidly for local or non-local firms, then investors' have the opportunity to act more quickly on information for local or non-local firms. For this analysis, I look at professional investor requests on EDGAR in either the two days and one hour following the release of public filings on EDGAR. Similar to equation (1), I evaluate the timing of information retrieval using the following research design:

$$Retreival\ Delay_{i,j,t} = \delta_1 \cdot Local_{i,j} + \delta_{i,t} + \delta_{j,t} + \mu_{i,j,t} \quad (2)$$

Where *Retrieval Delay* is measured as the difference in minutes between the investor's retrieval time of the filing and the time at which the filing was released. Since the exact time of day that

[16] I retain only one factor from this analysis based on tests of eigenvalues being greater than one, the scree plot, and a parallel analysis.

[17] I categorize investors as following a firm if they search for a firm's financial information in two or more calendar quarters. Conclusions are similar in sign, significance, and magnitude if observations with information acquisition of zero are excluded.

[18] Giannini et al. (2015) uses a cutoff of 100 miles, while Ivkovic and Weisbenner (2007) uses a cutoff of 250 miles. The results are not sensitive to these alternative cutoffs.

the filings were released is unknown, I use the timing of the first acquirer as a proxy. If δ_1 is significantly different from zero, then this suggests that investors value the timeliness of public disclosure access differently for local and non-local investments.

The second method I employ to test *H2* is a more direct measure of value and focuses on investors' trading profitably. Prior literature suggests that investors are more informed about a stock if their buy decisions are more positively correlated with future returns (Maffett, 2012). I expand this structure in two steps. First, I investigate whether buy decisions are more informed when coupled with public information acquisition. As such, I interact this relation with public information acquisition. Second, to test if public information is more valuable for local or non-local investors (*H2*), I further interact the relation with an indicator for local investments. The culminating research design is as follows:

$$\begin{aligned} Return_{j,t+1} = {} & \beta_1 \cdot Return_{j,t} + \beta_2 \cdot Return_{j,t-1} + \beta_3 \cdot Buyhold_{i,j,t} + \beta_4 \cdot Local_{i,j} \\ & + \beta_5 \cdot Search_{i,j,t} + \beta_6 \cdot Local_{i,j} \cdot Search_{i,j,t} + \beta_7 \cdot Local_{i,j} \cdot Buyhold_{i,j,t} \\ & + \beta_8 \cdot Search_{i,j,t} \cdot Buyhold_{i,j,t} + \beta_9 \cdot Search_{i,j,t} \cdot Local_{i,j} \cdot Buyhold_{i,j,t} \\ & + \beta_{i,t} + \epsilon_{j,t+1} \end{aligned} \quad (3)$$

Return is the quarterly firm stock return adjusted by the value-weighted market return. *Buyhold* is an indicator if the investor's quarter end shares were greater than or equal to the previous quarter. *Search* is an indicator if the institutional investor searched for firm information on EDGAR during the quarter.[19,20] The main coefficient of interest is β_9. If β_9 is positive and significant, then EDGAR research is more informative to local investors' portfolio decisions than

[19] I do not include firm-time fixed effects because of perfect correlation with the dependent variable. In robustness tests, I find that results are similar using a firm fixed effect.

[20] I use an indicator variable in place of prior proxies for information acquisition for ease of interpreting results. Conclusions are identical using other measures of information acquisition.

non-local investors. To test this, I manually match EDGAR professional investors to their 13(f) portfolio holdings on Thomson Reuters. I link EDGAR names to Thomson Reuters' names using two steps. First, I look through each candidate name match based on high name similarity. Second, I manually search through Thomson Reuters' names for common firms on EDGAR that do not match in the previous step.[21] This results in matching 609 institutions to the EDGAR research data. In instances where an institutional investor has multiple locations searching on EDGAR (i.e., multiple branches), I assign the most frequent location to that institutional investor.[22]

For the same sample of professional investors with 13(f) holdings, I investigate if "Local Bias" is related to the amount of public information acquired (*H3*):

$$PB_{i,j,t} = \gamma_1 \cdot Local_{i,j} + \gamma_2 \cdot PIA_{i,j,t} + \gamma_3 \cdot Local_{i,j} \cdot PIA_{i,j,t} + \alpha_{i,t} + \alpha_{j,t} + \varepsilon_{i,j,t} \quad (4)$$

Where PB is one of two measures of portfolio bias as of the end of the calendar quarter. The first measure, *PB1*, is the difference between the investor's portfolio weight in stock *i* and the market portfolio weight in stock *i*. *PB2* is the log of the investor's portfolio weight in stock *i* divided by the market portfolio weight in stock *i* (Cooper et al. 2012). If γ_2 and γ_3 are insignificantly different from zero, then γ_1 represents the previously documented "Local Bias" in portfolio holdings. If γ_3 is positive and significant, then "Local Bias" is larger when investors acquire public information, consistent with public information acting as a contributing force to "Local Bias" (Van Nieuwerburgh and Veldkamp, 2009).

The above specifications result in the formulation of three distinct samples. The first sample, which I label the *Primary Sample*, is associated with testing equation (1). The second

[21] I define high name similarity using SAS' compged function being less than 500. I define common firms as the top 200 most frequent firms in EDGAR that do not match based on high similarity.

[22] Results are robust to assigning the zip code location used on the investors' 13(f) filing.

sample, associated with testing equation (2), is labeled the *Delay Sample*. The third sample, associated with testing equations (3) and (4), is labeled the *13(f) Sample*.

CHAPTER 4: RESULTS AND ROBUSTNESS

4.1 Tests of Hypothesis 1 – Proximity and Public Information Acquisition

4.1.1 Main Analysis

Table 2 presents the descriptive statistics for the *Primary Sample*. The sample has 6,624,370 investor-firm quarters. In about 10 percent of these observations, the investor is within 100 kilometers of the firm. The mean number of kilobytes requested each quarter for a firm was about 535. For perspective, the median byte size for form 10-K in the sample is 225 kilobytes. The average number of files downloaded for the first time in a quarter by an EDGAR user for a specific company is 0.89, with 0.81 redownloaded files. The correlations presented in Panel B of Table 2 provide initial evidence that *Local* tends to be positively related to measures of information acquisition, suggesting that the closer an investor is to an investment, the more information they acquire about that investment.

Table 3 presents results from estimating equation (1) for each of the measures of public information acquisition. Consistent with the correlations, *Local* is positively related to public information acquisition. In other words, the same investor, with both a local and non-local investment, prefers to acquire more information about local investments. On average, investors tend to acquire 20.16% more kilobytes of information, 5.83% more previously unseen files, and redownload 7.36% more files if they are within 100 kilometers of the investment. Using the mean value of the total bytes acquired each quarter (*SumSize*) from Table 2, this suggests that being proximate to an investment results in about 107 additional kilobytes of information being acquired for a company each quarter. This is roughly equivalent to a single investor requesting

about 6 more material disclosures (i.e., 8-Ks) for one company every quarter or about 5 more annual reports (i.e., 10-Ks) over the lifetime of the investor-firm relationship.[23] These results are consistent with predictions that investors will acquire more public information for local investments (Van Nieuwerburgh and Veldkamp, 2009).[24] In untabulated analyses, I find that the research by local investors tends to be particularly concentrated around informational events, which is consistent with their increased research being information and trade-based.

To investigate how pervasive investors' preferences are to acquire more local information, I look at cross-sectional variation in investor behavior based on the filing type being requested. It is possible that proximity is only related to increased acquisition of certain company filings, such as annual or quarterly reports. I explore this possibility by looking at investors that acquired at least some information during the quarter. Table 4 presents estimates of Equation (1) using *InitialDownloads* for groups of filings provided to the SEC. Forms are categorized into one of the following seven form categories: (1) Annual Reports, (2) Quarterly Reports, (3) IPO Reports, (4) Material Disclosures, (5) Insider Trading Reports, (6) Proxy Statements, or (7) Other Forms.[25] The first noteworthy result is that acquisition of information for each type of information is positively related to an investor's proximity to investment. Interestingly, the relation between proximity and information acquisition is the smallest for annual and quarterly reports. The relation between proximity and public information acquisition seems to be strongest

[23] The median 8-K and 10-K byte size in the sample is about 18 and 225 kilobytes respectively.

[24] Van Nieuwerburgh and Veldkamp (2009) characterizes investors' information acquisition choice in their proposition two. In addition to predicting that investors will learn more about local investments, this proposition highlights that investors have incentives to learn about (1) larger risk factors, and (2) risk factors for which the average investor is uncertain. In untabulated analyses, I find evidence consistent with these additional predictions. Specifically, investors acquire more public information for larger and more volatile investments.

[25] The specific forms associated with each category is as follows: (1) 10-Ks, (2) 10-Qs, (3) S1s, (4) 8-Ks, (5) Forms 4, 5, 6, (6) DEF Forms, or (7) Other. For more detail on each form category, please refer to https://www.sec.gov/info/edgar/forms/edgform.pdf.

amongst material disclosures and insider trading reports. Being local to an investment is related to increased information acquisition of material disclosures and insider trading reports by 6.71% and 10.42%, respectively. This evidence suggests that the relation between proximity to investment and information acquisition permeates across information types, but that there is variation in the strength of the relation between information types. Additionally, if local investors do have an information advantage, it seems that it is likely to be derived from their increased use of material news and insider trading information.

I perform a second cross-sectional split based on the length of the investor-firm relationship in Table 5. One concern is that local investors' preference for additional public information might be transient and only exist at the onset of the investor-firm relationship. If asymmetric information acquisition does not persist, then this would be generally inconsistent with local investors having a consistent information advantage. In contrast to this concern, results in Table 5 suggest that investors' preferences are actually increasing throughout the investor-firm relationship. Across all four measures of information acquisition, I find that a longer investor-firm relationship increases investors' preference to acquire more public information for local stocks. This finding is consistent with the view that investors obtain private information through strong and established relationships and community ties to local firms (e.g., Coval and Moskowitz, 2001).

4.1.2 Exogenous Shock to Investor Proximity:

Deciding to use EDGAR as a source for firm information and selecting firms to follow are both choices made by investors. It is possible that professional investors using EDGAR are not a representative sample of the investor population or that other reasons lead an investor to both select firms and request information. Possible selection biases or correlated omitted

variables could result in observing a positive relation between investor proximity and information acquisition. To address these concerns, I exploit plausibly exogenous variation in investor proximity to a firm resulting from changes in firm headquarter location *after* investors select to follow a firm.

According to extant academic research (Strauss-Kahn and Vives, 2009), firms decide to move headquarter locations based on strategic incentives regarding lower taxes, lower employee wages, and better airport facilities. It seems reasonable to assume that any given investor's preference for proximity is immaterial in executives' decision criteria to move a firm's headquarters and the location selected. This assumption is reinforced by the fact that firm investors are geographically spread across the country, with each investor possibly vying for proximity to the firm.

To implement this research design, I identify changes in firm headquarter locations using changes in the zip code of a firm's business address listed on their annual and quarterly filings. I only select firms who change zip code locations once during my sample period and who move more than 100 kilometers away from their initial location. I further exclude firms for which the zip code switches back and forth between two locations. With this set of firms, I structure a simple difference-in-differences research design, where previously untreated (i.e., non-local) investors are arguably "randomly" treated with proximity.

Table 6 presents the findings of this difference-in-differences research design using the eight quarters before and after firms move headquarters. I further include investor-firm fixed effects, which controls for fixed biases and incentives that may have motivated an investor to follow and research about a firm. The combination of these fixed effects, along with firm-time fixed effects subsumes both difference-in-differences main effect indicators for the post period

(*Post*) and being treated with proximity (*Treat*). Table 6 shows that for each of the proxies for public information acquisition, being treated with proximity results in increased information acquisition, relative to professional investors who remained untreated. [26] Collectively, these tests add further evidence that, irrespective of endogenous selection, proximity to investment increases investors' likelihood of acquiring public information (*H1*).

4.2 Test of Hypothesis 2 - Value of Information Acquisition for Local vs Non-local Investors:

The prior analyses suggest that investors prefer local information and that their preference is consistent with them having a private information advantage in local stocks. I next test whether public information is more valuable to local or non-local investors using two approaches: (1) enhanced timeliness tests and (2) improved portfolio decision-making tests.

4.2.1 Enhanced Timeliness of Access:

Using the *Delay Sample,* I investigate whether investors acquire local or non-local information more rapidly. The average delay in information retrieval by professional investors over the two-day period after the release of the filing is 7.57 hours. The average delay in information retrieval over a 60-minute period is around 6.7 minutes. Table 7 presents results for investors that acquire information in a 48-hour period as well as a 60-minute period following the release of the filing. I find evidence that over a 48-hour period, the same investor arrives about 30 minutes more rapidly to local information compared to non-local information. Over a 60-minute period, I find that investors acquire local information about 1 minute more rapidly

[26] One concern is that changes in familiarity, rather than private information access, play a role in investors' information acquisition response. I discuss familiarity more generally in subsequent robustness tests. In this specific setting, familiarity concerns are mitigated for three reasons. First, investors already knew and searched for the firm before the change in headquarters and should therefore already be familiar with the firm. Second, changes in headquarters are typically announced well in advance of the actual move. As such, familiarity effects for "treated" investors should appear in the pre-headquarter change period, minimizing the likelihood of observing a post-treatment effect. Third, in untabulated analyses I find that the treatment effect is larger for investors that had followed the firm for a longer period of time. These investors are arguably highly familiar with the firm, and should have a reduced treatment effect if the treatment effect were capturing changes in familiarity.

than non-local information. For the average investor in the 60-minute window this time advantage equates to approximately a 12.7% reduction from the average. These results suggest that investors derive value from local information.

One concern with this analysis is that west coast investors might be asleep when east coast disclosures are made, or alternatively east coast investors are no longer working when west coast disclosures are made. If this is the case, then a time-zone effect might explain the above findings. The analysis focusing on the hour following the disclosure release helps mitigate this concern, but to more fully rule out a mechanical relation, I perform an additional robustness check. In untabulated analyses, I find similar results when only focusing on disclosures made between the hours of 11:00 am and 5:00 pm EST. These hours of disclosure represent the period of time where both east coast and west coast investors are awake and likely working. As such, the enhanced timeliness of local investors in acquiring newly released disclosures does not seem to be coming from a mechanical time-zone effect.

4.2.2 Improved Decision Making:

Using the *13(f) Sample*, I test whether the value of public information varies by investors' proximity to investment. Table 8 presents descriptive statistics on this sample. The median investment in a firm by institutional investors in this sample is $2.6 million. On average institutional investors perform EDGAR research for approximately 21% of their invested stocks. Roughly 8% of investments are local to institutional investors. Table 9 presents tests on improved decision-making. In columns (1) and (3) we see that research on EDGAR is associated with increased decision-making ability for institutional investors. Specifically, institutions' buy and sell decisions tend to be approximately 27 basis points more profitable when coupled with public information acquisition. In columns (2) and (4), we see that research on EDGAR for local

firms is incrementally more informative relative to non-local firms. We can interpret column (4) results as the same investor deriving more portfolio value from researching local stocks than non-local stocks, with the order of magnitude being approximately 54 basis points. The combination of these results and the timeliness results suggest that investors value public information more for local investments.

4.3 Test of Hypothesis 3 – "Local Bias" and Public Information Acquisition:

I next test whether "Local Bias" is related to the amount of public information acquired. Table 10 helps us better understand the forces involved. The first noteworthy result in Table 10 is that, across both proxies for portfolio bias, investors exhibit "Local Bias." At the individual stock level, local stocks tend to be 0.069% larger in investors' portfolios relative to the market portfolio. Secondly, we see that investors' "Local Bias" is increasing in the amount of public information acquired, consistent with public information being a contributing force to "Local Bias." In terms of magnitude, a one standard deviation increase in public information acquisition (*Factor*) is associated with a 0.054% increase in "Local Bias." In untabulated robustness tests, I find that altering the empirical design yields similar results. Specifically, including the lagged level of bias produces similar results. Additionally, including an investor-firm fixed effect, which forces variation to exist intertemporally for the investor-firm pair, yields similar results. These results corroborate the theory in Van Nieuwerburgh and Veldkamp (2009) which predicts that public information acquisition plays a contributing role in investors' "Local Bias."

4.4 Robustness of Empirical Results:

4.4.1 Investment Level Control:

One concern with some of the preceding analyses is whether the results are being driven by a correlated omitted variable related to the degree of investment in the company. In other

words, investors research more for stocks that make up a larger portion of their portfolio. To address this concern, I use the *13(f) Sample* and estimate equations (1), (2), (3), and (4) including the level of investment as an additional control variable. Table 11 presents the coefficients of interest from each of the equations after including the level of past investment. Corroborating the initial concern, I find that the amount previously invested by the institutional investor is positively related to public information acquisition. Of greater importance for this paper, investors' preference to acquire local information is still positive and significant (*H1*). Furthermore, controlling for the amount previously invested does not change inferences regarding local information being more valuable for portfolio decision (*H2*). Lastly, the relation between portfolio bias and public information acquisition is unaffected (*H3*).

4.4.2 Familiarity Bias:

Familiarity bias is a behavioral bias of wanting to own investments with which an investor is familiar (Huberman, 2001). Importantly, it does not represent an information advantage in local stocks. A potential concern is whether familiarity bias or a prior private information advantage is the driving force behind the documented findings. In an attempt to address this concern, I highlight five points. First, the data represents research by an investor for a firm and therefore suggests that the investor is already familiar with the investment opportunity being researched (i.e., they know it exists). Second, to the extent that familiarity is a firm fixed effect, similar to visibility, the research design controls for this type of variation. Third, in looking at the first quarter an investor researched on EDGAR for a firm, arguably a time of similar familiarity for local and non-local investments, I find that investors still research more for local investments. Fourth, prior robustness tests that control for past investment should simultaneously control for the degree of familiarity since research suggests "familiarity breeds

investment." Fifth, since familiarity bias is unrelated to information advantages it would not explain why public information for local stocks is more valuable to investors' portfolio decisions. Collectively, these arguments suggest the evidence in this paper is more consistent with prior private information advantages motivating investors to acquire additional public information.

4.4.2 Controlling for Relative Wealth Hedging:

Some of the documented findings are consistent with relative wealth hedging as well as asymmetric information advantages.[27] Using a model, Garcia and Strobl (2011) concludes if agents have incentives to "catch up with the Joneses," then agents will herd on acquiring the same information. To the extent that local stocks hedge relative wealth concerns, this theory suggests that investors will acquire more information about proximate investments. If this were the driving force behind the documented asymmetric public information acquisition, then there should be an increased incentive for investors to acquire local information if neighboring investors also acquire local firm information (i.e., herding on information). In untabulated analyses, I find that herding is unrelated to the documented local effect. Specifically, investors' preference to acquire more public information for local investments is not increasing in the amount of other proximate investors searching for the same firm.

4.4.3 Information Asymmetry:

The main analyses find that investors acquire more public information for local investments and that public information tends to be more valuable to local investors. In combination, this suggests that having additional local investors requesting public information should expand firm-level information asymmetries. I look at firm-level information asymmetry

[27] Evidence surrounding relative wealth effects contributing to home bias are mixed. For example, Massa and Simonov (2006) conclude that home bias is not explained by nonfinancial income.

on public disclosures dates to confirm this line of reasoning. Using disclosure dates on EDGAR, I evaluate if having local investors access the disclosed filing increases firm-level information asymmetry. Since firms often make multiple disclosures in a calendar quarter, I use a firm-time fixed effect to control for fixed firm-level information environment attributes. Table 12 presents the findings of this analysis. In columns (1) through (4), I find that information asymmetry is higher when local investors access the public filing. Column (2) includes stock turnover and stock return to control for attributes specific to the news event that may attract more local investors. Columns (3) and (4) include disclosure type fixed effects to further control for the nature of the news being disclosed. Across each of these specifications, I find consistent evidence that having local investors access a disclosed filing tends to be associated with increased firm-level information asymmetry.[28]

4.4.4 Non-Professional Investors:

I explore the generalizability of results to non-professional investors. It's possible that other investor groups behave contrary to the consistently documented behavior. For example, retail investors might have investment objectives or risk tolerances that are different from professional investors. These differences could alter preferences to acquire more public information for local or non-local firms. It is also possible that non-professional investors tend to receive a different amount or type of private local information, which may also manifest in changes in information acquisition behavior. I attempt to address this element of generalizability through the use of non-professional investors' search on EDGAR. Non-professional investors are all EDGAR users who do not own the IP address from which they are searching, and instead

[28] These findings are generally robust to changes in measurement and research design. Results hold (1) using the percentage of investors that are local instead of an indicator variable for the presence of local investors, (2) including a control for the number of investors accessing the filing, (3) relaxing the fixed effects structure, and (4) using an alternative measure of information asymmetry based on bid-ask spread.

receive an IP address through an Internet Service Provider. These investors are more likely to be retail investors, but could also capture smaller-sized professional investors or even non-investor financial statement users. I use this sample of investors in untabulated analyses. I find that these EDGAR users continue to exhibit a preference to acquire more public information for local investments.

4.4.5 Alternative Measure of Proximity:

The entire set of analyses pivot around the geodesic proximity between investors and investments (i.e. as a crow flies). This focus is important given the literature on "Local Bias." I step back from this focus temporarily to explore whether information acquisition behavior responds similarly to variation in the travel time between investors and firms. This measure is based on the time it takes to travel between two locations and is an indicator set to one if the duration of the optimal travel itinerary is less than or equal to two hours (*Short Travel)*. This measure comes from prior literature that uses estimates of travel time by car and plane between locations and selects the optimal travel pattern in a given period of time (Giroud, 2013; Bernstein et al., 2016).[29] In addition to proximity, this measure also captures elements of economic distance. Travel time between two locations encompasses the economic cost of monitoring a firm as well as the cost of maintaining firm relationships. Table 12 presents empirical findings for equation (1) as it relates to this measure of proximity. Similar to prior results, public information acquisition is increasing in proximity.

[29] Refer to the Appendix for additional details on the construction and measurement of this variable.

CHAPTER 5: CONCLUSION

It is often posited that investors have a private information advantage in local stocks, which causes them to overweight local stocks in their investment portfolio (i.e., "Local Bias"). I investigate the role of *public* information for local and non-local investor. This paper asks the following three questions: First, do investors mitigate their information disadvantage by acquiring more public information for non-local stocks (i.e., level the playing field) or do they maximize their information advantage by acquiring more public information for local stocks (i.e., widen the gap)? Second, does the value of public information vary by an investor's proximity to investment? Third, does public information consumption mitigate or moderate investors' "Local Bias?"

The body of evidence in this study suggests that investors use public information to widen their local information advantage. Using a novel dataset of professional investors' requests on EDGAR, a repository for all publicly filed financial disclosures with the SEC, I find that investors choose to acquire more information for local investments. For example, each quarter investors access 20.16% more kilobytes of information if they are within 100 kilometers of the investment. For the average investor in this sample, being proximate to an investment is roughly equivalent to requesting about 6 additional material disclosures (i.e., 8-K's) for that company in one quarter. These inferences are robust to variations in research design, sample selection, and measurement. For a sample of non-local investors, I show that when investors are exogenously "treated" with proximity because of changes in firm headquarter location, they increase their public information consumption.

I next show that public information is more valuable to local investors using two approaches. First, I show that in both the two days and one hour following the release of public information, investors tend to acquire local information more rapidly than non-local information. This enhanced speed may afford investors the opportunity to act on the information more rapidly and suggests that investors value the timeliness of public information more for local than non-local investments. As a second test of value, I show that local investors' share purchase decisions are more positively correlated with future returns when coupled with public information acquisition. In tandem, these findings suggest that public information is more valuable to local investors.

I lastly test whether "Local Bias" in stock ownership varies with the amount of public information acquired. If public information acquisition is a contributing force, then additional public information should enhance investors' "Local Bias." Consistent with prior literature, I show that the institutional investors in my sample exhibit "Local Bias." More importantly, I find that investors' "Local Bias" is increasing in the amount of public information acquired about local firms. This evidence corroborates prior findings that "Local Bias" is related to information asymmetries, and further suggests that public information plays a contributing role.

Admittedly, these analyses are subject to important caveats. While the documented findings are quite robust, it is possible that these relations are not generalizable to investors who do not use EDGAR as a source for firm information. Additionally, omitted variable bias concerns may persist, so long as those explanations are unrelated to investor or firm attributes. While acknowledging these concerns and caveats, this paper provides consistent results that proximity to investment is positively related to information retrieval, that public information

tends to be more valuable to local investors, and that public information enhances investors' "Local Bias."

In closing, I anticipate the identification and use of professional investor requests for financial information to be valuable to future research initiatives. I expect future work investigating information acquisition at the investor level to be very fruitful given the combination of granular investor data on information retrieval and the importance of understanding the value of information disclosed.

APPENDIX 1: VARIABLE DEFINITIONS AND MEASUREMENT

Variable	Description
Abn_Return	The quarterly firm stock return adjusted by the value-weighted market return.
Buyhold	An indicator if the institutional investor's quarter end shares were greater than or equal to the previous quarter end shares.
Factor	The principal component of *InitialDownloads* , *SumSize* , and *Redownloads* .
Illiquidity	The natural log of the Amihud price impact of trade measured on the release date of the EDGAR filing (Amihud, 2002). The price impact of trade is calculated as the absolute value of the firm stock return divided by the volume of dollars traded.
InitialDownloads	The number of files downloaded for the first time by that investor.
Local	An indicator if the geodesic distance in miles between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 miles.
Local Investors	An indicator for the EDGAR filing being accessed by a local professional investor on the release date.
PB1	The difference between investor portfolio weight in stock i and the market portfolio weight for stock i times 100.
PB2	The log of the investor portfolio weight in stock i divided by the market portfolio weight in stock i .
Redownloads	The number of files redownloaded (i.e. requested for the second time or more).
Relationship Age	Length of monitoring, in quarters, of the firm by the investor.
Retrieval Delay	The difference in minutes between the investors retrieval time of the filing and the time at which the filing was released. Since the exact time of day in which the filing was released is unknown, the time of the first viewer is used as a proxy.
Return	The firm stock return on the release date of the EDGAR filing.
Search	An indicator if the institutional investor searched for firm information on EDGAR during the quarter.
Short Travel	An indicator if the number of minutes needed to optimally travel between the investor's IP address location and the firm's headquarters (Giroud, 2013) is less than two hours.
SumSize	The aggregate kilobyte size for all files downloaded for the first time.
Turnover	The natural log of the volume of shares traded divided by the shares outstanding on the release date of the EDGAR filing.

Calculation of Short Travel:

I preform similar steps to Giroud (2013) to obtain the optimal travel itinerary between two locations. The following 5 steps describe this process. (1) Obtain flight routes and their respective travel duration data from the T-100 Domestic Segment Database. (2) Construct a database of optimal flight itineraries between any two locations, with a maximum number of 5 flights in sequence. (3) Obtain estimates of travel time by car via API requests from Bing Maps. These requests include travel time estimates between investors and companies as well as travel time to their nearby airports.[30] (4) Create a database of possible travel plans for each investor-firm combination (i.e. driving only vs. driving to a nearby airport and flying to an airport nearby the firm).[31] (5) Identify the fastest travel itinerary for the investor-firm combination.[32]

[30] Airport zip code location is obtained from the FAA: http://www.faa.gov/airports/airport_safety/airportdata_5010/
[31] Nearby airports are defined as the five closest airports to the zip code and all airports within 200 miles.
[32] Following Giroud (2013), 60 minutes is added to the flight itineraries for checking in baggage. Additionally, 60 minutes is added for each layover.

REFERENCES

Ahearne, A. G., W. L. Griever, and F. E. Warnock. 2004. Information costs and home bias: An analysis of US holdings of foreign equities. *Journal of International Economics* 62 (2): 313-36.

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

Bernile, G., A. Kumar, and J. Sulaeman. 2015. Home away from home: Geography of information and local investors. *Review of Financial Studies* 28 (7): 2009-49.

Bernile, G., A. Kumar, J. Sulaeman, and Q. Wang, 2016. Has local Information Advantage Disappeared? Working paper.

Bernstein, S., X. Giroud, and R. R. Townsend. 2016. The impact of venture capital monitoring. *The Journal of Finance* 71 (4): 1591-1622.

Mullins, B., Rothfeld, M., McGinty, T., and Strasburg, J. 2013. Traders pay for an early peek at key data. *Wall Street Journal* June 13.

Bozanic, Z., J. Hoopes, J. Thornock, B. Williams. 2017. IRS Attention. *Journal of Accounting Research* 55 (1): 79-114.

Choe, H., B. Kho, and R. Stulz. 2005. "Do domestic investors have an edge? The trading experience of foreign investors in Korea." *The Review of Financial Studies* 18 (3): 795-829.

Cohen, L. 2009. Loyalty-based portfolio choice. *Review of Financial Studies* 22 (3): 1213-45.

Cooper, I., P. Sercu, R. Vanpée. 2012. The Home Equity Bias Puzzle: A Survey. *Foundations and Trends in Finance* 7 (4): 289-416.

Coval, J. D., and T. J. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance* 54 (6): 2045-73.

Coval, J. D., and T. J. Moskowitz. 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 109 (4): 811-41.

Drake, M. S., P. Quinn, and J. R. Thornock. 2017. Who Uses the Financial Statements? A Demographic Analysis of Financial Statement Downloads from EDGAR. *Accounting Horizons* 31 (3): 55-68.

Drake, M. S., D. T. Roulstone, and J. R. Thornock. 2015. The determinants and consequences of information acquisition via EDGAR. *Contemporary Accounting Research* 32 (3): 1128-61.

Drake, M. S., D. T. Roulstone, and J. R. Thornock. 2016. The usefulness of historical accounting reports. *Journal of Accounting and Economics* 61 (2-3): 448-464.

Ferri, R., S. Garrett, F. Holmes, G. McBride, T. Odean, G. Papadopoulos, C. Rotblut, G. Sauter, M. Thakor, L. Zimpleman. 2013. The experts: How much should you invest abroad? 2013. *Wall Street Journal (Online)*, Jun 11, 2013.

French, K. R., and J. M. Poterba. 1991. Investor diversification and international equity markets. *The American Economic Review* 81 (2): 222-6.

Garcia, D., and G. Strobl. 2011. Relative wealth concerns and complementarities in information acquisition. *Review of Financial Studies* 24 (1): 169-207.

Giannini, R., P. Irvine, T. Shu. 2015. Do Local Investors Know More? A Direct Examination of Individual Investors' Information Set. Working Paper.

Giroud, X. 2013. Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics* 128 (2) (May 01): 861-915.

Gormley, T. A., and D. A. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27 (2): 617-61.

Graham, J. R., C. R. Harvey, and H. Huang. 2009. Investor competence, trading frequency, and home bias. *Management Science* 55 (7): 1094-106.

Gow, I., D. Taylor, and R. Verrecchia. 2012. On The Complementary Relation between Earnings and Private Information. Working Paper.

Huang, Y., H. Qiu, Z. Wu. 2012. Local Bias in Investor Attention: Evidence from China's Internet Stock Message Boards. Working Paper.

Huberman, G. 2001. Familiarity breeds investment. *Review of Financial Studies* 14 (3): 659-80.

Investment Company Institute. 2015. Investment Company Fact Book. Washington, DC.

Ivković, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance* 60 (1): 267-306.

Lee, C., B. Mucklow, and M. Ready. 1993. "Spreads, depths, and the impact of earnings information: An intraday analysis." *The Review of Financial Studies* 6 (2): 345-374.

Lee, C., P. Ma, and C. Wang. 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116 (2): 410-31.

Lewis, K. 1999. "Trying to explain home bias in equities and consumption." *Journal of Economic Literature* 37 (2): 571-608.

Loughran, T., and B. McDonald. 2015. Information decay and financial disclosures.

Massa, M., and A. Simonov. 2006. Hedging, familiarity and portfolio choice. *Review of Financial Studies* 19 (2): 633-85.

Maffett, Mark. 2012. "Financial reporting opacity and informed trading by international institutional investors." *Journal of Accounting and Economics* 54 (2): 201-220.

Mondria, J., T. Wu, and Y. Zhang. 2010. The determinants of international investment and attention allocation: Using internet search query data. *Journal of International Economics* 82 (1): 85-95.

Morse, A., and S. Shive. 2011. Patriotism in your portfolio. *Journal of Financial Markets* 14 (2): 411-40.

Seasholes, M. S., and N. Zhu. 2010. Individual investors and local bias. *The Journal of Finance* 65 (5): 1987-2010.

Securities and Exchange Commission (SEC). 2000. Final Rule: Selective Disclosure and Insider Trading. Available at: https://www.sec.gov/rules/final/33-7881.htm#P12_1307. SEC Offices, Washington D.C.

Securities and Exchange Commission (SEC). 2010. Important Information About EDGAR. Available at: https://www.sec.gov/edgar/aboutedgar.htm. SEC Offices, Washington D.C.

Strauss-Kahn, V., and X. Vives. 2009. "Why and where do headquarters move?" *Regional Science and Urban Economics* 39 (2): 168-186.

Thomson Reuters. 2008. User's Guide to Thomson Reuters Mutual Fund and Investment Company Common Stock Holdings Databases on WRDS.

Van Nieuwerburgh, S., and L. Veldkamp. 2009. Information immobility and the home bias puzzle. The Journal of Finance 64 (3): 1187-215.

Figure 1A: United States Zip Codes with Professional Investors using EDGAR

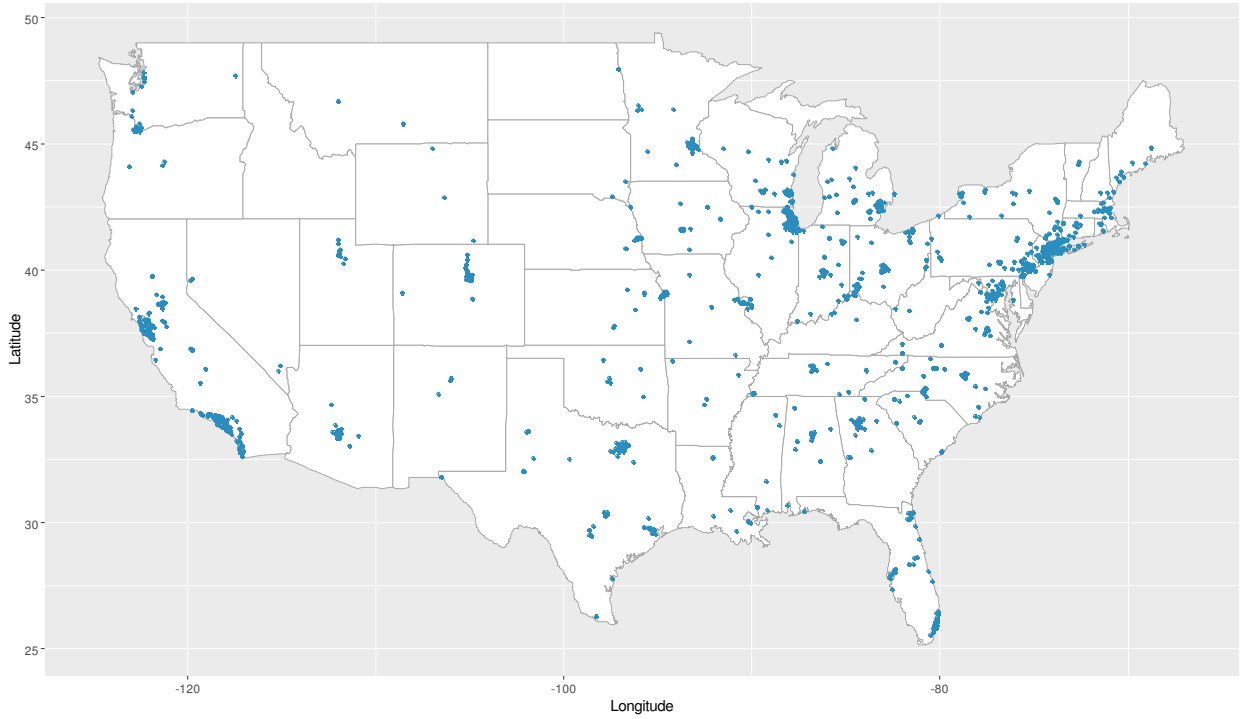

Figure 1B: United States Zip codes with Firm Headquarters

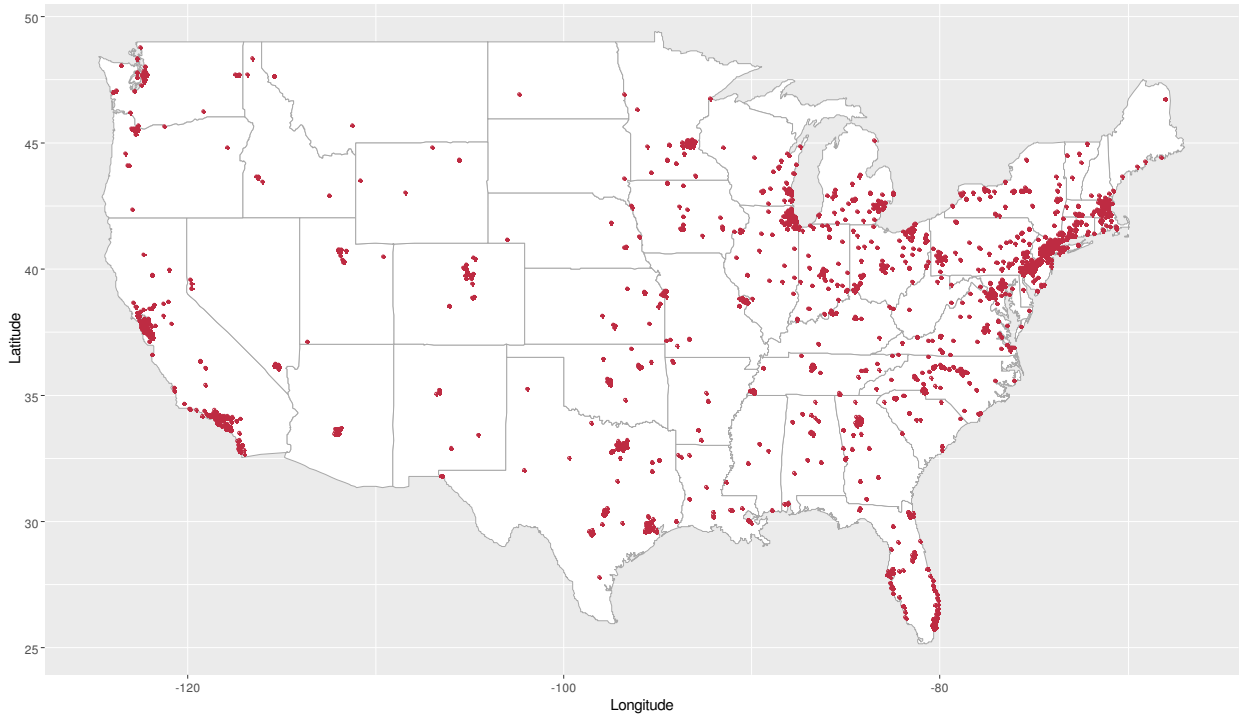

Table 1. Professional Investor Descriptive Statistics

Panel A. Top 10 Most Research Active Professional Investor Firms

	Investor Firm	Average Research in a Quarter
1	JP MORGAN	21,570.10
2	CITIGRP	18,067.04
3	UBS	11,648.66
4	BANK OF AMERICA	9,023.53
5	WACHOVIA	7,922.88
6	MORGAN STANLEY	7,360.39
7	MERRILL LYNCH	7,122.00
8	DEUTSCHE BANK	6,530.31
9	WELLS FARGO	6,509.14
10	DEAN WITTER	6,296.57

Panel B: Top 10 Most Frequent Cities for Professional Investor Firms

	City, State	Number of Firms
1	New York, NY	446
2	Atlanta, GA	143
3	Chicago, IL	73
4	Boston, MA	51
5	San Francisco, CA	48
6	Los Angeles, CA	39
7	Greenwich, CT	33
8	Minneapolis, MN	27
9	Washington, DC	27
10	Jersey City, NJ	25

Panel C: Time Series Patterns in the Number of Investors Using EDGAR

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Number of Investors	1,720	2,714	2,687	2,330	2,580	2,910	3,171	3,048	2,938	3,019	2,844	2,732	1,926

Notes: This table presents the most common professional investor firms using EDGAR (Panel A), the most common cities for professional investor firms using EDGAR (Panel B), and the time-series pattern in the number of professional investors using EDGAR (Panel C). *Average Research in a Quarter* is the average number of total filings requested by the professional investor firm each quarter. Professional investor firms are made up of several professional investors.

Table 2. Descriptive Statistics of *Primary Sample*

Panel A. Summary Statistics

	Q1	Median	Mean	Q3	Std Dev.	Obs
Local	0.00	0.00	0.10	0.00	0.30	6,624,370
SumSize	0.00	0.00	534.60	258.22	1,488.20	6,624,370
InitialDownloads	0.00	0.00	0.89	1.00	1.87	6,624,370
Redownloads	0.00	0.00	0.81	0.00	3.16	6,624,370
Factor	-1.09	-1.09	0.00	1.01	1.52	6,624,370
Relationship Age	3.00	7.00	8.84	12.00	7.64	6,624,370

Panel B. Correlation Statistics

		(1)	(2)	(3)	(4)	(5)	(7)
Local	(1)		0.017	0.018	0.020	0.025	-0.002
SumSize	(2)	0.020		0.952	0.489	0.603	-0.100
InitialDownloads	(3)	0.030	0.531		0.546	0.823	-0.059
Redownloads	(4)	0.029	0.320	0.532		0.590	0.015
Factor	(5)	0.025	0.603	0.823	0.590		-0.105
Relationship Age	(7)	-0.002	-0.100	-0.059	0.015	-0.105	

Notes: This table presents descriptive statistics on proximity and public information acquisition using quarterly data on investor-company pairs. The sample starts in the first quarter of 2003 and ends in the first quarter of 2015. Panel A presents summary statistics. Panel B presents Pearson (below diagonal) and Spearman (above diagonal) correlation statistics. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads,* and *Redownloads.* Variables are not logged to preserve interpretability. *Relationship Age* is the length of monitoring, in quarters, of the firm by the investor.

Table 3. Relation between Public Information Acquisition and Investors' Proximity to Investment

	SumSize			*InitialDownloads*			*Redownloads*			*Factor*		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Local	0.184	7.629	0.000	0.057	7.126	0.000	0.071	8.625	0.000	0.154	7.964	0.000
Investor x Year-qtr FE	Yes			Yes			Yes			Yes		
Firm x Year-qtr FE	Yes			Yes			Yes			Yes		
R-squared	0.255			0.225			0.176			0.245		
Obs	6,624,370			6,624,370			6,624,370			6,624,370		

Notes: This table presents the relation between public information acquisition and investors' proximity to investment. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownloads.* All continuous variables are logged to reduce the influence of outliers. Standard errors are clustered by Investor and Firm.

Table 4. Relation between Public Information Acquisition and Investors' Proximity to Investment by Information Type Requested

		InitialDownloads				
	Form Type	coeff	t-stat	p-value	R-squared	Obs
Local	*Annual Reports*	0.010	4.919	0.000	0.167	906,207
	Quarterly Reports	0.012	5.499	0.000	0.154	1,012,412
	IPO Reports	0.018	4.002	0.000	0.306	198,771
	Material Disclosures	0.065	6.836	0.000	0.208	676,763
	Insider Trading Reports	0.099	6.227	0.000	0.544	62,459
	Proxy Statements	0.014	4.394	0.000	0.360	212,265
	Other Forms	0.089	3.995	0.000	0.261	643,233
	Investor x Year-qtr FE				Yes	
	Firm x Year-qtr FE				Yes	

Notes: This table presents cross-sectional results of how the filing type requested relates to an investors inclination to research more for local stocks. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *InitialDownloads* is the number of files downloaded for the first time by that investor. All continuous variables are logged to reduce the influence of outliers. Standard errors are clustered by Investor and Firm.

Table 5. Cross-Sectional Variation in the Relation between Public Information Acquisition and Investors' Proximity to Investment: Relationship Age

	SumSize			*InitialDownloads*			*Redownloads*			*Factor*		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Local	0.135	6.880	0.000	0.039	6.770	0.000	0.038	5.458	0.000	0.098	6.711	0.000
Relationship Age	-0.036	-18.981	0.000	-0.005	-14.363	0.000	0.002	5.798	0.000	-0.010	-10.952	0.000
Local x Relationship Age	0.009	4.260	0.000	0.003	4.596	0.000	0.004	5.415	0.000	0.008	4.951	0.000
Investor x Year-qtr FE	Yes			Yes			Yes			Yes		
Firm x Year-qtr FE	Yes			Yes			Yes			Yes		
R-squared	0.258			0.227			0.177			0.246		
Obs	6,624,370			6,624,370			6,624,370			6,624,370		

Notes: This table presents cross-sectional results of how the strength of the investor-firm relationship relates to an investor's inclination to research more for local stocks. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *InitialDownloads* is the number of files downloaded for the first time. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownload. Relationship Age* is the length of monitoring, in quarters, of the firm by the investor, where the first quarter takes a value of zero. Standard errors are clustered by Investor and Firm.

Table 6. Relation between Public Information Acquisition and Investors' Proximity to Investment Using Changes in Firm Headquarters

	SumSize			*InitialDownloads*			*Redownloads*			*Factor*		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Post x Treat	0.133	2.063	0.039	0.038	2.775	0.006	0.034	2.214	0.027	0.094	2.624	0.009
Investor x Firm FE		Yes			Yes			Yes			Yes	
Firm x Year-qtr FE		Yes			Yes			Yes			Yes	
R-squared		0.279			0.295			0.315			0.322	
Obs		476,135			476,135			476,135			476,135	

Notes: This table compares investors' public information acquisition surrounding exogenous changes in proximity to investment using a simple difference-in-differences setup. Each of the included investors are initially non-local, with certain investors becoming local as a result of changes in firm headquarter locations. *Post x Treat* equals one if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownload.* Standard errors are clustered by Investor and Firm.

Table 7. Relation between the Delay in Public Information Retrieval and Investors' Proximity to Investment

	Retrieval Delay					
	48 Hours			*1 Hour*		
	coeff	t-stat	p-value	coeff	t-stat	p-value
Local	-29.488	-3.981	0.000	-0.852	-4.974	0.000
Investor x Year-qtr FE	Yes			Yes		
Firm x Year-qtr FE	Yes			Yes		
R-squared	0.242			0.310		
Obs	975,034			410,884		

Notes: This table evaluates whether investors acquire public disclosures more rapidly based on their proximity to investment. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *Retrieval Delay* is the difference in minutes between the investor's retrieval time of the filing and the time at which the filing was released. Since the exact time of day in which the filing was released is unknown, the time of the first viewer is used as a proxy. All standard errors are clustered by Investor and Firm.

Table 8. Descriptive Statistics for the *13(f) Sample*

	Q1	Median	Mean	Q3	Std Dev.	Obs
Local	0.00	0.00	0.08	0.00	0.27	2,120,687
Search	0.00	0.00	0.21	0.00	0.41	2,120,687
Buyhold	0.00	1.00	0.57	1.00	0.50	1,918,475
Abn_Return	-0.09	-0.01	0.00	0.08	0.15	1,913,998
PB1	-0.04	0.00	0.26	0.03	0.03	1,946,397
PB2	-1.79	-0.47	-0.62	0.62	2.27	1,946,385
Investment ($ Millions)	0.44	2.57	53.30	15.28	343.69	1,913,998

Notes: This table presents descriptive statistics using quarterly data on investor-company pairs for the sample of institutional investors with identifiable search data on EDGAR. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *Search* is an indicator if the institutional investor searched for firm information on EDGAR during the quarter. *Buyhold* is an indicator if the institutional investor's quarter end shares were greater than or equal to the previous quarter. *Abn_Return* is the quarterly firm stock return adjusted by the value-weighted market return. *PB1* is the difference between investor's portfolio weight in stock i and the market portfolio weight for stock i times 100. *PB2* is the log of the investor's portfolio weight in stock i divided by the market portfolio weight in stock i. *Investment* is the dollar value the institutional investor has invested in the given firm.

Table 9. Public Information Acquisition and Informed Trading

	Abn_Return $_{t+1}$											
	(1)			(2)			(3)			(4)		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Buyhold $_t$	-0.001	-2.002	0.045	-0.001	-1.980	0.048	-0.002	-2.394	0.017	-0.002	-2.350	0.019
Search $_t$	0.001	0.616	0.538	0.001	0.881	0.379	0.001	0.666	0.506	0.001	0.891	0.374
Buyhold $_t$ *x Search* $_t$	0.003	2.310	0.021	0.002	1.824	0.068	0.003	2.921	0.004	0.002	2.718	0.007
Local $_t$				-0.001	-0.694	0.488				-0.001	-0.630	0.529
Buyhold $_t$ *x Local* $_t$				0.000	0.124	0.902				0.000	0.118	0.906
Search $_t$ *x Local* $_t$				-0.003	-1.518	0.129				-0.002	-1.067	0.286
Buyhold $_t$ *x Search* $_t$ *x Local* $_t$				0.007	2.903	0.004				0.005	2.446	0.015
Abn_Return $_t$	-0.042	-4.478	0.000	-0.042	-4.476	0.000	-0.047	-5.077	0.000	-0.047	-5.076	0.000
Abn_Return $_{t-1}$	-0.062	-9.615	0.000	-0.062	-9.616	0.000	-0.030	-4.760	0.000	-0.030	-4.761	0.000
Investor x Year-qtr FE	No			No			Yes			Yes		
R-squared	0.008			0.008			0.052			0.052		
Obs	1,557,520			1,557,520			1,556,741			1,556,741		

Notes: This table presents the value of public information for portfolio decisions, where a higher correlation between purchases and future returns suggests increased value. *Abn_Return* is the quarterly firm stock return adjusted by the value-weighted market return. *Buyhold* is an indicator if the institutional investor's quarter end shares were greater than or equal to the previous quarter. *Search* is an indicator if the institutional investor searched for firm information on EDGAR during the quarter. *Local* is an indicator if the the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. Standard errors are clustered by Investor and Firm.

Table 10. Relation between Local Bias and Public Information Acquistion

	PB1						*PB2*					
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Local	0.069	2.254	0.025	0.069	2.113	0.035	0.136	2.457	0.014	0.114	2.264	0.024
Factor				0.045	5.365	0.000				0.184	13.660	0.000
Local x Factor				0.054	2.232	0.026				0.073	3.206	0.001
Investor x Year-qtr FE	Yes			Yes			Yes			Yes		
Firm x Year-qtr FE	Yes			Yes			Yes			Yes		
R-squared	0.472			0.472			0.412			0.421		
Obs	1,945,684			1,945,684			1,945,672			1,945,672		

Notes: This table presents the relation between local bias and public information acquisition. *PB1* and *PB2* are measures of portfolio bias. *PB1* is the difference between the investor's portfolio weight in stock *i* and the market portfolio weight in stock *i*. *PB2* is the log of the investor's portfolio weight in stock *i* divided by the market portfolio weight in stock *i*. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownloads.* Standard errors are clustered by Investor and Firm.

Table 11. Main Findings Controlling for the Level of Investment

Equation	Dependent Variable	Independent Variable	coeff	t-stat	p-value	R-squared	Obs
(1)	*SumSize*	*Local*	0.348	4.988	0.000	0.442	1,977,931
(1)	*InitialDownloads*	*Local*	0.098	4.578	0.000	0.483	1,977,931
(1)	*Redownloads*	*Local*	0.097	4.398	0.000	0.360	1,977,931
(1)	*Factor*	*Local*	0.184	5.033	0.000	0.469	1,977,931
(2)	*Retrieval Delay (Two-days)*	*Local*	-21.782	-2.483	0.013	0.257	291,886
(2)	*Retrieval Delay (One-hour)*	*Local*	-0.687	-2.341	0.019	0.376	109,296
(3)	Abn_Return_{t+1}	$Buyhold_t$ x $Search_t$ x $Local_t$	0.005	2.400	0.017	0.052	1,556,741
(4)	*PB1*	*Local x Factor*	0.044	1.848	0.065	0.502	1,738,423
(4)	*PB2*	*Local x Factor*	0.013	2.800	0.005	0.795	1,738,413

Notes: This table presents the coefficients of interest from the main analyses controlling for the level of investment held by the institutional investor in the stock at the beginning of the quarter. *Local* is an indicator if the geodesic distance between the investor's IP address location and the firm's headquarters (i.e. as the crow flies) is less than or equal to 100 kilometers. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownloads. Retrieval Delay* is the difference in minutes between the investor's retrieval time of the filing and the time at which the filing was released. Since the exact time of day in which the filing was released is unknown, the time of the first viewer is used as a proxy. *Abn_Return* is the quarterly firm stock return adjusted by the value-weighted market return. *Buyhold* is an indicator if the institutional investor's quarter end shares were greater than or equal to the previous quarter. *Search* is an indicator if the institutional investor searched for firm information on EDGAR during the quarter. *PB1* is the difference between the investor's portfolio weight in stock i and the market portfolio weight in stock i. *PB2* is the log of the investor's portfolio weight in stock i divided by the market portfolio weight in stock i.

Table 12. Firm-Level Information Asymmetry and Local Research

	Illiquidity											
	(1)			(2)			(3)			(4)		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Local Investors	0.027	2.534	0.015	0.049	4.155	0.000	0.021	1.936	0.059	0.039	3.498	0.001
Turnover				-0.332	-16.276	0.000				-0.338	-16.346	0.000
Return				1.609	4.255	0.000				1.612	4.275	0.000
Firm x Year-qtr FE	Yes			Yes			Yes			Yes		
Event Type FE	No			No			Yes			Yes		
R-squared	0.915			0.919			0.916			0.920		
Obs	223,745			223,745			223,745			223,745		

Notes: This table presents the relation between firm-level information asymmetry and the presence of local investors accessing the disclosed EDGAR filing on the filing release date. *Illiquidity* is the natural log of the Amihud price impact of trade measured on the release date of the EDGAR filing (Amihud, 2002). The price impact of trade is calculated as the absolute value of the firm stock return divided by the volume of dollars traded. *Local Investors* is an indicator for the EDGAR filing being accessed by a local professional investor on the release date. *Turnover* is the natural log of the volume of shares traded divided by the shares outstanding on the release date of the EDGAR filing. *Return* is the firm stock return on the release date of the EDGAR filing. Event Type FE is a fixed effect for the seven categories of forms filed with the SEC: (1) Annual Reports, (2) Quarterly Reports, (3) IPO Reports, (4) Material Disclosures, (5) Insider Trading Reports, (6) Proxy Statements, or (7) Other Forms. Standard errors are cluster by Firm and Year-qtr.

Table 13. Relation between Public Information Acquisition and Investors' Proximity to Investment Using Alternative Measure of Proximity

	SumSize			*InitialDownloads*			*Redownloads*			*Factor*		
	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value	coeff	t-stat	p-value
Short Travel	0.150	6.962	0.000	0.043	6.283	0.000	0.052	7.458	0.000	0.118	7.025	0.000
Investor x Year-qtr FE	Yes			Yes			Yes			Yes		
Firm x Year-qtr FE	Yes			Yes			Yes			Yes		
R-squared	0.254			0.220			0.172			0.240		
Obs	5,941,184			5,941,184			5,941,184			5,941,184		

Notes: This table presents the relation between public information retreival and an investor's travel time to firm headquarters. *Short Travel* is an indicator if the number of minutes needed to optimally travel between the investor's IP address location and the firm's headquarters is less than or equal to two hours. *SumSize* is the aggregate kilobyte size for all files downloaded for the first time. *InitialDownloads* is the number of files downloaded for the first time. *Redownloads* represents the number of files redownloaded. *Factor* is the principal component of *SumSize, InitialDownloads, and Redownloads.* All continuous variables are logged to reduce the influence of outliers. Standard errors are clustered by Investor and Firm.