The Geographic Dispersion of Google Search and the Market Reaction to Earnings Announcements[*]

Sabrina Chi
Devin Shanthikumar[†]

2015 AFA Annual Meeting
This Draft: December 19, 2014
First Draft: November 2012

Abstract

We examine the impact of distance on investor search behavior, and the effect of geographic dispersion of investor search on the stock market response around earnings announcements. We find significant "local bias" in Internet search behavior. While more visible firms have more geographically dispersed search, there is significant additional variation in search dispersion. Motivated by theories of network effects and psychological distance, we predict and find that firms with a higher geographic dispersion of search experience higher abnormal trading volume, lower abnormal bid-ask spreads, and larger earnings response coefficients at the time of earnings announcements, as well as weaker post-earnings-announcement drift. These results hold both cross-sectionally and when examining changes in dispersion or propensity-score matched subsamples. In addition, path analysis suggests that both network effects and investor psychology are significant drivers of the return results. Overall, our results suggest that geographic proximity affects search, and that firms with more geographically dispersed search experience better market responses to earnings announcements.

Keywords: geography, Google, investor attention, information asymmetry, earnings response coefficient, post-earnings-announcement drift, investor psychology, network effects, local bias

---

[*] We thank Cyndall Hill, Chris Hines, Josh Hunt, YiFan Li, Stacey Kaden, Chris Knickerbocker, Jaclyn Prentice, Marc Rios, Timothy Seidel and Lili Zhou for excellent research assistance. We thank Eugene Soltes for graciously sharing press coverage data, Terrance Odean for sharing brokerage account data, and Bo Lu for sharing the nonbipartite matching algorithm. We thank two anonymous referees, Philip Bromiley, David Hirshleifer, Chad Larson, Alastair Lawrence, Scott Liao, Linda Myers, James Myers, Morton Pincus, Terry Shevlin, Siew Hong Teoh, Nicole Thibodeau, Mohan Venkatachalam, Charles Wang, Yong Yu and workshop participants at The University of Arkansas, The University of California Irvine, the American Accounting Association 2013 Western Regional Meeting, the American Accounting Association 2013 Annual Meeting, the 2014 AAA Financial Accounting and Reporting Section Annual Meeting, the 2014 NTU/NCCU Accounting Symposium, and the 2014 Southern California Research Forum for helpful comments.
[†] University of California, Irvine, CA 92617. Send correspondence to dshanthi@uci.edu.

1. *Introduction*

The ways in which investors obtain information about a firm have changed dramatically over the last twenty years. *Newsweek* devoted a February 1995 issue to articles about the burgeoning Internet. One editorial was titled, "The Internet? Bah. (Computers cannot replace books, teachers or newspapers)" (Stoll 1995). In 2013, Newsweek stopped publishing a paper edition and moved entirely to digital formats (Brown and Shetty 2012, www.newsweek.com January 2013). Today, an interested investor can quickly, easily, and inexpensively access much of a firm's news and information (e.g., SEC filings, press releases, and analyst earnings forecasts) through websites like Yahoo! and Google Finance. In this paper, we suggest that while the Internet has lowered information acquisition costs substantially, geography is still important in the Internet era. In particular, we suggest that investors are more likely to be aware of and interested in firms located nearby, so that even though they are able to search for any distant firm, they will choose to search for local firms more often. We also predict that the breadth of investor interest, measured by the average distance of investors searching for information about a firm, will impact how information is incorporated into the firm's stock price. We find evidence supporting these predictions.

A growing body of research examines the dynamics and effects of information dissemination and finds that press coverage reduces information asymmetry, increases investor response to information, and reduces mispricing of information (e.g., Bushee, Core, Guay and Hamm 2010; Soltes 2010; Engelberg and Parsons 2011; Drake, Guest and Twedt 2014). However, for the market to react to information, it is not sufficient that the information is published broadly. Hirshleifer, Lim and Teoh (2009), and Drake, Roulstone and Thornock (2012) find evidence suggesting that investors need to pay attention to, or demand, information in order for it to have an impact. We focus on a specific dimension of investor information demand: geography. We use state-level search information from Google's Search Volume Index (SVI) for searches within the United States, to create a measure of the geographic dispersion of investors searching for a given firm, for the period 2005 through 2011 (see Section 3.1 for details). Prior literature (e.g., Coval and Moskowitz 1999; Ivkovic and Weisbenner 2005) has established that investors have a preference for owning and trading stock of firms headquartered nearby (referred to as "local bias"). We predict that investors searching for a firm will disproportionately be those that are located near the firms' headquarters, but that certain firm characteristics will increase non-local investors' interest and thus the geographic dispersion of search. Consistent with our expectations of a local bias effect in Internet search, we find that for 81% of firm-years in our sample, the firm's headquarters state has a higher level of search than expected. In addition, average firm-searcher distances are similar to firm-investor distances documented in prior literature on local bias (Ivkovik and Weisbenner 2005). We also find that geographically dispersed search is positively related to expected firm characteristics such as retail industry, S&P 500 index membership, and a larger number of shareholders.

Given the local bias in Internet search behavior, we ask whether the geographic dispersion of search impacts the market's reaction to information. Two streams of literature suggest a link: first, literature on network effects and information dissemination, and second, literature in psychology.

First, prior work suggests that investors spread information by word of mouth in local networks (Hong, Kubik and Stein 2004 and 2005). There may be limited, or no, information about a firm within certain local networks, particularly those located far from a firm, if only investors close to a firm search for information. In contrast, if the same number of investors search, but those investors are spread around the country, the information will reach more networks and investors in those diverse networks can obtain the information through within-network word-of-mouth communication. Thus, broader geographic dispersion of search would lead to more investors being informed. In addition, prior literature provides evidence that broader *information dissemination* is associated with reduced information asymmetry around earnings announcements and faster incorporation of information into price (e.g., Loughran and Shultz 2005; Bushee, Core, Guay and Hamm 2010), as are higher investor attention and information demand (e.g., Merton 1987; Hirshleifer, Lim and Teoh 2009; Drake, Roulstone and Thornock 2012). Together, these past results suggest that broader *geographic dispersion of search* is similarly associated with information asymmetry and the incorporation of information into price, through the mechanism of increased investor trading.

Second, the Construal-Level Theory of Psychological Distance (Trope and Liberman 2010) demonstrates that responses to physically distant objects involve more abstract thinking, comparisons, and long-term thinking than responses to nearby objects. These thought processes will (arguably) lead to a smarter and more rational response to information like an earnings announcement. Thus, the theory suggests that broader geographic dispersion of search will improve the incorporation of information into price, through the mechanism of smarter investor trading.

We test whether the geographic dispersion of information demand impacts the market's response to earnings announcements. We control for other factors that affect the dissemination of information around earnings announcements, such as the overall abnormal level of search and the number of news articles about a firm. First, we examine whether liquidity and information asymmetry are related to geographic dispersion of search, as predicted by the theory of network effects. We find that higher geographic dispersion is associated with higher abnormal trading volume and lower abnormal bid-ask spreads around the earnings announcement. Second, we examine whether earnings information is incorporated into prices more quickly, as predicted by both network theory and psychological theory. We find that higher geographic dispersion of search is related to a stronger return response to earnings surprises during the announcement window, as well as lower subsequent post-earnings-announcement drift, suggesting that more dispersed investor interest improves the market response to earnings information. Our results suggest that the breadth of geographic interest in a firm, as measured by the geographic dispersion of Google search for a firm, is related to a distinct improvement in the market response to earnings announcements. To better capture whether geographic dispersion of search impacts the market response, we examine firms before and after changes in dispersion. We find that the firms with the highest increases in geographic dispersion of search experience a statistically and economically significant improvement in market responses to earnings, both in absolute terms and relative to firms with the largest decreases in dispersion. For example, firms in the top tercile by change of dispersion experience roughly an 80-95% increase in abnormal volume around earnings announcements and a 70-85% drop in abnormal spreads. In a similar spirit, we examine pairs of firms matched based on firm characteristics including firm visibility measures, but differing in geographic dispersion, using propensity-score matching. Our matching generates subsamples of firms which are well-matched along most firm and visibility characteristics. Yet the treatment firms, with higher geographic dispersion of search, experience significantly lower abnormal spreads, higher trading volumes, higher earnings response coefficients, and lower post-earnings-announcement drifts, than the matched control firms with lower geographic dispersion of search. The magnitudes of the effects are large, for example the reduction in post-earnings-announcement drift amounts to 25-88%, depending on the exact specification used. Finally, path analysis shows that each of the two mechanisms predicted by network and psychological theories – increased investor trading and smarter investor trading, respectively – contributes to our results. Trading volume explains roughly 20-25% of the total effect of geographic dispersion of search on earnings response coefficients and post-earnings-announcement drift, while our proxy for smart investor trading explains 12-22% of the total effect.

We contribute to the literatures on local bias, information dissemination, and investor attention. Prior research has shown that geography matters to investors in their investing choice and their responses to newspaper articles (Coval and Moskowitz 1999; Ivkovic and Weisbenner 2005; Engelberg and Parsons 2011; Miller and Shanthikumar 2012). We show that geography continues to play a significant role in investors' behavior during the Internet era, despite the potentially significant drop in the costs of learning about distant firms. Prior research has also shown that investor attention and information demand are important for the market's response to earnings information (Hirshleifer, Lim and Teoh 2009; Drake, Roulstone and Thornock 2012). We show that geography is an important dimension of investor interest. The geographic breadth of interest improves the market's reaction to earnings information, due to both network effects in information diffusion, which lead to information reaching more investors, and the psychology of distance, which leads to distant investors making smarter and more rational trading decisions. These results contribute to our understanding of information diffusion, investor attention, and the effects of information demand on the market's response to information.

The remainder of the paper is organized as follows. Section 2 discusses relevant prior literature and our predictions. Section 3 describes the data. We present our empirical methodology and results in Section 4. Finally, Section 5 concludes.

2. *Prior Literature and Empirical Predictions*

A growing literature has shown evidence that broader information dissemination is related to lower information asymmetry (e.g., lower bid-ask spreads, greater depths, higher trading volume) (Bushee, Core, Guay and Hamm 2010; Soltes 2010) and better pricing of accounting information (Drake, Guest and Twedt 2014). However, even if information is disseminated, investors need to pay attention to, or demand, information in order for it to have an impact. For example, Hirshleifer, Lim and Teoh (2009) find weaker responses to earnings announcements and stronger post-earnings-announcement drift when more firms announce earnings on the same day (a proxy of the level of investor distraction). Drake, Roulstone and Thornock (2012) use Google search behavior as a proxy for investors' information demand and find that when there are more searches for a firm prior to an earnings announcement, prices reflect more earnings information in the pre-announcement window. If investors demand information, prices are more likely to reflect that information.

Building on this literature, we aim to better understand investors' demand for information and the impact that information demand has on the overall market response to firm-specific information. The specific dimension that we focus on is geography, motivated by a large literature on the effects of proximity on investors and information intermediaries.

The role of geography in investor interest

Prior literature has established that investors have a preference for locally headquartered firms (Coval and Moskowitz 1999; Ivkovic and Weisbenner 2005; Seasholes and Zhu 2010). The preference for local stocks may be due to familiarity, local information advantage, and stronger personal wealth ties such as employment. Even in the Internet era, those factors are likely to drive investor interest. We predict that investors today continue to have a greater interest in local firms, and that this interest is displayed in Internet search. In addition, previously documented local effects vary due to firms' visibility. For example, local effects appear to be weaker for larger firms (Coval and Moskowitz 1999) and S&P 500 firms (Ivkovic and Weisbenner 2005). Thus, we also predict that search is more geographically dispersed for firms that are larger and more visible. Formally, our first hypotheses, stated in alternative form, are as follows:

H1a: Investors search disproportionately for firms located close to them.

H1b: Investors' searches for distant firms are positively associated with firms' visibility.

While we predict that investors display "local bias" in search behavior, it is important to note that certain factors may diminish local bias for Internet search, when compared to the "local bias" of ownership and trading documented in prior literature (e.g., Ivkovic and Weisbenner 2005). Considering local newspapers, for example, investors across the country can now access a firm's local paper online. In the pre-Internet-era local investors responded much more strongly when local press covered a firm than non-locals did (Engelberg and Parsons 2011; Miller and Shanthikumar 2012), but this may no longer be the case. Further, investors across the country can access the firm's earnings announcement online before any newspaper reporting. Thus, local bias may not be significant in the Internet era, or at least may not translate to Internet search activity. Collectively, it is an empirical question whether local bias affects Internet activity.

The role of geography in the market response to information

Prior literature has shown that broader information dissemination is related to lower information asymmetry around earnings announcements (Bushee, Core, Guay and Hamm 2010) and more complete pricing of accounting information (Drake, Guest and Twedt 2014) and that higher investor attention is related to stronger event-window return responses to earnings and lower post-earnings announcement drift (Hirshleifer, Lim and Teoh 2009). The dimension that we add to these studies is the geography of investors' information search. Based upon two theories, the theory of network effects in the spread of information and Construal-Level Theory of Psychological Distance, we hypothesize that the geographic dispersion of search is positively related to the market response to information, in terms of lower information asymmetry and stronger responses to the information, just as broader information dissemination and higher investor attention are. Specifically, we predict that greater geographic dispersion of search will be associated with lower bid-ask spreads, higher trading volume, a stronger return response to earnings announcements and lower post-earnings-announcement drift.[1]

Theories of social networks suggest that if information reaches a larger number of

[1] Studies of local bias suggest that local investors and analysts are better informed than non-locals (e.g., Coval and Moskowitz 2001; Ivkovic and Weisbenner 2005; Malloy 2005; Engelberg and Parsons 2011; Miller and Shanthikumar 2012; and in contrast, Seasholes and Zhu 2010). Ayers, Ramalingegowda and Yeung (2011) find results consistent with local institutional investors with large holdings playing a stronger monitoring role than distant ones, again consistent with local investors being better informed. However this literature does not lead to any specific predictions for *the informational efficiency of stock prices*. More local trading may lead to the incorporation of more information in stock prices, or it may lead to greater information asymmetry. The abnormal returns earned by local investors suggests that prices may be less informationally efficient for firms with more local bias, thus providing more opportunities for informed investors to profit.

networks, for example through more geographically dispersed investor interest, more investors will acquire and potentially trade on the information. Hong, Kubik and Stein (2004) develop a model for the influence of social networks on investors and find empirical evidence consistent with social network effects. Hong, Kubik and Stein (2005) show that mutual fund managers are more likely to trade a particular stock if other fund managers in the same city are trading the stock, and to trade in the same direction. Together, this evidence is consistent with investors spreading information by word of mouth. We refer to this as the *network theory*. Geographically dispersed investors searching for a firm are likely to facilitate information flow within many local networks across the country, increasing the number of investors who ultimately receive the information.[2] Because individuals pass information through word of mouth, if one investor in a network searches for information on Google, the other individuals in his network are more likely to learn the information even if they do not conduct Google searches themselves. Thus, if 100 individuals from 100 different networks learn about a news event from Google searches, that information is likely to reach individuals in 100 networks. On the other hand, if 100 individuals from one network learn about the news event from Google searches, that information is likely to reach the individuals in only one network. Thus, the theory and evidence of network effects suggests that broader geographic dispersion of search leads to more investors (investors in more networks) receiving and responding to information, with associated decreases in information asymmetry and increases in the market reaction for information events. Our second set of tests focuses on earnings announcements as an important information event. We predict that

[2] In a similar vein, Loughran and Schultz (2005) provide evidence that urban firms attract more analyst following and institutional investors than rural firms, suggesting that the ease of information access is greater for urban firms. Stock trading volume is higher and bid-ask spreads are lower for urban firms, consistent with information diffusion being greater in urban areas, and with that information diffusion improving liquidity.

firms with a higher geographic dispersion of search experience lower information asymmetry around the earnings announcement, captured by higher abnormal trading volume and abnormal bid-ask spreads, and a stronger reaction to earnings information, captured by higher earnings-response-coefficients at the time of the announcement and lower post-earnings-announcement drift in the subsequent months.

The Construal-Level Theory of Psychological Distance suggests that individuals' thought and decision processes are influenced by the psychological distance they perceive between themselves and a particular object, such as a firm, and that psychological distance and physical distance tend to be related (Trope and Liberman 2010). The longer the distance between an individual and an object, the less the individual focuses on details and the more the individual thinks abstractly about the object and focuses on "central" characteristics. In our setting, a distant investor is likely to focus on more "central" characteristics of the firm, such as prior-year performance, CEO turnover or the information content of earnings news. In contrast, a local investor is likely to focus on details of the firm, such as construction of a new employee parking lot or donations to local schools, which are less relevant when evaluating the stock price implications of an earnings announcement. In addition, as greater psychological distance facilitates abstract thinking, a distant investor potentially can better compare firms with each other. Collectively, the Construal-Level Theory of Psychological Distance suggests that greater physical distance facilitates a smarter and more rational investor response to information about a firm. We refer to this as the *psychological theory*.[3]

[3] The application of Construal-Level Theory of Psychological Distance to any particular situation is difficult given the complexity and abstractness of the theory. Regardless of which interpretation of the theory is correct, Construal-Level Theory of Psychological Distance predicts that distance impacts investors' judgments. It is an empirical question as to whether distance helps or hinders reactions to earnings announcements.

If distant investors respond to news more rationally, then a more geographically dispersed investor set will improve the incorporation of information into price. Thus psychological theory also predicts a positive relation between the geographic dispersion of search and price responses around earnings announcements.

Based on *network theory*, we predict H2 as follows:

> H2: Broader geographic dispersion of search reduces information asymmetry around earnings announcements.

Based on both *network* and *psychological theories*, we predict H3 as follows:

> H3: Broader geographic dispersion of search increases the return response to earnings information around earnings announcements, and reduces post-earnings-announcement drift.

Finally, because both *network theory* and *psychological theory* lead to the prediction of H3, we test the theories more directly by examining the mechanism underlying the relation between geographic dispersion of search and the price response to earnings, using path analysis. We explain the path analysis in more detail in Section 4.3.

3. Data, Variable Measurement and Research Design

The sample consists of stocks listed on NYSE, AMEX and NASDAQ from 2005 through 2011, with CRSP and Compustat data, and with state-level annual Google search data. To identify firms in Google Trends, we use ticker symbols, as in Da, Engelberg and

---

To expand upon the potential alternative interpretation: one could use Construal-Level Theory to argue that greater physical distance impairs, rather than improves, an investors' response to information. If one believes that small details are vital to the optimal investor response to information, then local investors have an advantage since they pay more attention to small details. In addition, distant investors may underestimate small probabilities, since they tend to disregard unlikely events. If small probabilities are important to a proper reaction, then local investors have an advantage. We argue that the abilities to focus on central characteristics, ignore small details, and compare firms more abstractly will on average be more important for interpreting earnings announcements than the abilities to recall small details and properly weight small probabilities, however we acknowledge this potential alternative application of the theory.

Gao (2011). There are 1,529 distinct tickers in the initial sample. We remove tickers with alternate meanings, such as "LAKE", "MAIN" and "RENT," require non-missing data for key variables, and eliminate penny stocks. The final sample contains 945 distinct firms and 21,597 firm-quarter observations. Table 1, Panel A outlines the sample selection process.

3.1 Measuring the Geographic Dispersion of Google Search

We collect state search data from Google Trends (http://www.google.com/trends), which tracks Google users' search volume by search term. Google's servers maintain a log of users' Internet Protocol (IP) addresses, which Google uses to identify the location of a computer which is used for a search. Google aggregates search data for each state and the District of Columbia, and then identifies the state with the most searches for a given term (the top state). It defines the search volume index (SVI) for each state as the ratio of searches from that state to searches from the top state, scaling the index to 100% for the top state.

Since Google does not provide the exact location of searches, we assume that all searches in a state originate from its geographic center. For a given firm, we weight each state-specific SVI by the distance between the firm's headquarters (from Compustat historical) and the geographic center of the search state (from 2010 Census), and take the average, excluding Hawaii and Alaska to avoid skewing our measure (Ivkovich and Weisbenner 2005). Thus our main geographic dispersion measure, *HQDisGD,* is defined as

$$HQDisGD = \frac{\sum_{X=1}^{49}(SVI\ for\ state\ X)*(distance_{firm\ headquarters,\ state\ X})}{\sum_{x=1}^{49}(SVI\ for\ state\ X)}, \quad (1)$$

where $distance_{firm\ headquarters,\ state\ X}$ is the distance between the firm's headquarters and the geographic center of state X. *HQDisGD* captures the average distance between an investor searching for the firm and the firm's headquarters.

3.2 Earnings Surprises

We compute unexpected earnings as $UE_{jq} = AE_{jq} - FE_{jq}$, where AE_{jq} is the announced quarterly earnings per share (EPS) of firm j in quarter q, and FE_{jq} is expected earnings. We use two pairs of announced and expected earnings: EPS before extraordinary items for the given quarter and for the prior year's same quarter, and the "actual" value of earnings from IBES along with the consensus analysts' earnings forecasts calculated from IBES detail forecast data (e.g., Bernard and Thomas 1990; Livnat and Mendenhall 2006; Hirshleifer, Lim and Teoh 2009).[4] The consensus analyst forecast is defined as the median of analysts' final forecasts over 60 trading days before the earnings announcement, with at least three analysts covering the firm. We scale UE_{jq} by price-per-share for firm j at the end of quarter q, preceding the announcement, following prior literature (e.g., Livnat and Mendenhall 2006; Hirshleifer, Lim and Teoh 2009), to calculate standardized unexpected earnings, *SUE*.

3.3 Announcement Period Responses and Post-Announcement Returns

We examine four aspects of market response: announcement period abnormal bid-ask spreads, abnormal trading volume, abnormal return, and post earnings announcement drift. Announcement period abnormal spreads, *AbSpreads[0,1]*, are calculated as average daily bid-ask spreads over the two-day period around the earnings announcement minus the average daily bid-ask spreads over trading days [-41, -11], where daily spreads are the difference between the quoted offer and bid prices, divided by the midpoint of the offer and

[4] Prior literature has shown that small, individual, investors are more likely to use a random-walk-based earnings expectation model, and react naively to earnings announcements, while institutional investors react in a more sophisticated manner, and analyst forecasts are more representative of their expectations (e.g., Lee 1992; Bhattacharya 2001; Ke and Petroni 2004; Battalio and Mendenhall 2005; Hirshleifer, Myers, Myers and Teoh 2008; Shanthikumar 2012). This results in two slightly different post-earnings-announcement drifts – one for each type of earnings expectation model (Ayers, Li and Yeung 2011). We use both models to ensure that our results are not driven by using the earnings-expectations model of only one type of investor, particularly given that individuals may be more likely to use Google search (Da, Engelberg and Gao 2011), and thus drive our geographic dispersion measure.

bid prices, multiplied by 100 (e.g., Bushee, Core, Guay and Hamm 2010; Soltes 2010). Following Hirshleifer, Lim and Teoh (2009), we compute firm abnormal trading volume, *AbVol[0,1]*, as the average of daily trading volume over the two-day period around the earnings announcement minus the average daily trading volume over trading days [-41, -11], where daily trading volume is the log of dollar trading volume, calculated using the product of the closing price and the number of shares traded. We compute the abnormal stock return, *CAR[0,1]*, as the sum of the abnormal return over days 0 and 1, where abnormal returns are the difference between the raw return from CRSP and the return on a portfolio of firms matched on size and book-to-market ratio.[5] Finally, to measure post-earnings-announcement drift, we use a 60 trading day window, similar to prior literature (e.g., Livnat and Mendenhall 2006; Hirshleifer, Lim and Teoh 2009; Ayers, Li and Yeung 2011). Bernard and Thomas (1989) show that the majority of drift occurs in the first 60 trading days after the announcement. We define *CAR[2, 61]* as the size- and book-to-market-adjusted cumulative abnormal returns for the 60-trading-day period [2, 61], relative to the earnings announcement date.

3.4 Control Variables

We include control variables associated with initial market reactions to earnings news and post-earnings-announcement drift, based on prior research. Specifically, firm size and book-to-market (Collins and Kothari 1989), institutional ownership (Teoh and Wong 1993), and analyst following (Shores 1990), have been shown to affect market reactions to earnings news. We also include the number of earnings announcements made by other firms on the

[5] The portfolios, constructed at the end of June each year, are the intersections of five portfolios formed on size and five portfolios formed on book-to-market, following Fama and French (1993). We thank Ken French for providing portfolio data (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

same day and share turnover based on Hirshleifer, Lim and Teoh (2009). Those factors also affect subsequent market responses to earnings surprises (Bernard and Thomas 1989; Bartov, Krinsky and Radhakrishnan 2000) so we include them in our analysis of post-earnings announcement drift. In addition, we include a firm's press coverage, as measured by Soltes (2010), to control for the general news environment around the earnings announcement.[6] For the abnormal trading volume analysis, we additionally control for market-wide variation in trading volume. For the abnormal bid-ask spreads analysis, we also control for the reciprocal of share price, which is a proxy for trading costs, and the number of employees and the number of shareholders, following Loughran and Schultz (2005) and Bushee, Core, Guay and Hamm (2009). Appendix A provides a summary of variable definitions and data sources.

3.5 Descriptive Statistics

Table 1, Panels B, C and D, report sample characteristics. Observations are evenly distributed from 2005 through 2011. 13.22% of sample firms are headquartered in California, followed by Texas (11.40%) and New York (8.64%). In addition, sample firms are mainly in the retail (13.00%), business services (10.66%), and electronic equipment (5.21%) industries.

[Insert Table 1]

Table 2 presents summary statistics. All financial statement variables are winsorized at the 1st and 99th percentiles. We take logs for all search-related variables. Our sample is weighted towards large firms (median market capitalization of $1,582 million), with a median of 6 analysts following the firm and 55.3% institutional ownership.

[Insert Table 2]

[6] We thank Eugene Soltes for providing press coverage data.

4. *Empirical Results*

Section 4.1 reports results related to search behavior. Section 4.2 examines the market response to earnings announcements. We discuss additional analyses in Section 4.3.

4.1 Geographic Dispersion of Google Search and its Related Factors

We examine whether location plays a role in Internet search activity. While we expect investors to be interested predominantly in local firms, whether "local bias" translates to Internet search activity is unclear. Thus this is an important first step in our analysis.

We first examine the average distance of an individual searching for a firm from the firm's headquarters, using the simplifying assumption that searches from a particular state originate from the center of that state. We find that on average, an investor searching for a firm is located 887 miles from the firm's headquarters. Ivkovik and Weisbenner (2005, Table 1) report that the average distance between an investor and the headquarters of firms in their portfolio is 917 miles, using retail investor data for 1991-1996. This suggests that "local bias" in search activity for our recent 2005-2011 period is similar in magnitude to local bias in stock ownership documented in prior literature for earlier time periods.[7]

We also compare the geographic distribution of search activity with the expected randomly distributed search activity. We use Census data to determine the location of internet users.[8] If internet users search for companies without regard to distance, then 13% of

---

[7] In untabulated analyses, we further verify that *HQDisGD* is related to local bias from prior literature by comparing our variable to ownership-based local bias measures using brokerage account data from the 1991-1996 period. Our untabulated results show a positive significant 14% correlation between geographic dispersion of ownership in the earlier period and *Ln(HQDisGD)*, for firms which survive the entire window. They also show that both variables have similar determinants. We thank Terrance Odean for providing the retail brokerage account data. This data is described in more detail in Barber and Odean (2000).

[8] We use data from the October 2007 Current Population Survey (CPS), "School Enrollment and Internet Use." We use survey answers for Internet usage, combined with census bureau population weighting variables, to calculate the percentage of US Internet-using households which are located in each state.

our sample firms would have their home states as the top search state. Instead, we find that 20% of our firms have their home states as the top search state, meaning that investors search for local firms with a higher frequency than expected. For 81% of firm-years, the firm's headquarters state is ranked lower (has more search) than expected based on our bootstrap analysis. This is significantly higher than a random 50% frequency, with $p<0.001$.

Overall, the evidence supports H1a that investors search disproportionately for local firms. We next examine variation in the geographic dispersion of search to test H1b. We predict that firm characteristics which increase visibility of the firm, such as size, advertising expenditures, and press coverage, are related to a higher level of geographic dispersion of search, i.e., to less local bias in search. We estimate the following model:

$$Ln(HQDisGD) = \beta_0 + \beta_1\, Ln(SVI) + \beta_2 \#of\ News + \beta_3 Urban + \beta_4 Ln\,(Size) + \beta_5 Adv\ Exp + \beta_6 Ln(EMP) + \beta_7 Ln(SHR) + \beta_8 Ln(AF) + \beta_9 IO + \beta_{10} SP500 + \beta_{11} Retail + \beta_{12} BM + \varepsilon, \qquad (2)$$

at the firm-year level. The variables are computed on an annual basis and are defined below. We include *Ln(SVI)*, the mean of weekly log Google search volume index for the firm, to examine the relation between *HQDisGD* and the overall level of search activity. The overall level of search may be a summary measure for a firm's visibility, and thus we would expect the two to be positively related. *# of News* is the number of articles in the *Wall Street Journal*, the *New York Times*, *USA Today*, and the *Washington Post* that mention the firm, which can help to increase visibility to investors (Bushee, Core, Guay and Hamm 2010; Soltes 2010). *Urban* is an indicator variable which takes value 1 if the firm is located in one of the 10 most populous cities and 0 otherwise. Loughran and Schultz (2005) show that firms in urban locations have higher trading volume, analyst coverage and institutional ownership than rural firms, suggesting that they are more visible to investors. *Ln(Size)* is the log of market equity value. Coval and Moskowitz (1999) show that local bias is weaker for larger firms, suggesting that those firms are more visible to distant investors. We include *Adv Exp*, advertising expense scaled by sales, since advertising expenditure is related to investor awareness (Grullon, Kanatas and Weston 2004; Lou 2014). We also include the number of employees and the number of shareholders, since firms with more of either may be more visible to investors (e.g., Hong, Kubik and Stein 2008; Bushee, Core, Guay and Hamm 2010). *Ln(EMP)* (*Ln(SHR))* is the log of 1 plus the number of employees (shareholders). *Ln(AF)* is the log of 1 plus the mean of analyst following; *IO* is the mean of shares owned by institutional investors scaled by total shares outstanding. Both higher analyst following and higher institutional ownership are associated with higher firm visibility (e.g., Bushee and Miller 2012). *SP500* is an indicator variable which takes value 1 if the firm is in the S&P 500 index. Ivkovik and Weisbenner (2005) find that local bias is weaker for S&P 500 firms. We suggest that retail firms may be more visible to the average individual, since they are consumer-facing. *Retail* is an indicator which takes the value 1 if the firm is in the retail, consumer goods or entertainment industry and 0 otherwise. Finally, we include the book-to-market ratio *BM*, book value of common equity divided by size, under the idea that low book-to-market firms may be "glamour" stocks in favor with investors (Lakonishok, Shleifer and Vishny 1994). For all regressions onwards, we run pooled OLS regressions and estimate standard errors with two-dimensional clustering (Petersen 2009; Gow, Ormazabal and Taylor 2010) by period (year for Table 3, quarter for Tables 4-10) and firm, and include state fixed effects to address the possibility of demographic differences unrelated to *HQDisGD.*

We present the results in Table 3. Columns 1 and 2 use the full sample, while Columns 3 and 4 restrict the sample to firm-years for which we have press coverage data. We use industry fixed effects in columns 2 and 4, instead of the dummy for *Retail*, to ensure that the results are robust to controlling for other industries that may have different national versus local visibility. As expected, we find that almost all of our visibility variables are related to more dispersed Google search (higher *HQDisGD).* Higher search levels (*Ln(SVI))*, newspaper coverage (*# of News)*, urban firms, larger firms, firms with more employees, shareholders, analyst following and institutional ownership, S&P 500 firms, and retail firms, are all associated with significantly more dispersed Google search. The results for these variables, associated with increased firm visibility, are significant at the 10% level or better in all four models using two-tailed tests, and in many cases the level of statistical significance is much higher. Lower book-to-market firms ("glamour" stocks) have significantly more dispersed Google search when we use the full sample (columns 1 and 2). The only visibility variable for which we do not find significant results is advertising expenditures.

While there is clearly a strong and statistically significant relation between the geographic dispersion of search and firm visibility, it is important to note that firm visibility does not fully explain the geographic dispersion of search. The adjusted R^2 values reported in Table 3 range from 7.6% to 11.6%. In untabulated analyses, we examine the correlations between *Ln(HQDisGD)* and the firm visibility measures in equation (2). None of the correlations is higher than 13.5%. To control for the higher dissemination of information that may occur for more visible firms, we include search levels and press coverage as control variables in all of our remaining tests. We also conduct a robustness test, presented in Section 4.3.3 and Table 10, in which we use the residual from Table 3 as our primary variable.

[Insert Table 3]

Together, these results suggest that "local bias" continues to exist in the Internet Era, and that it applies to investors' Internet search activity. This occurs even though distance does not affect the cost to search, as it does for other information-gathering activities.

4.2 Geographic Dispersion of Search and the Market Reactions to Earnings News

To test hypotheses H2 and H3, we focus on the incremental effect of (lagged) geographic dispersion, *HQDisGD*, on information asymmetry around earnings announcements (*AbSpreads[0,1]* and *AbVol[0,1]*) and the incorporation of earnings information into prices (*CAR[0,1]* and *CAR[2,61]*).

4.2.1 Information Asymmetry around Earnings Announcements: Spreads and Volume

We first examine the relation between the geography of search and information asymmetry. We regress *AbSpreads[0,1]*, our first information asymmetry measure, on the quarterly decile ranks of absolute value of earnings surprise (*R_absSUE*), the quarterly decile ranks of lagged *HQDisGD* (*R_HQDisGD*), each normalized to range from -0.5 to +0.5, as well as a set of control variables. This specification is similar to that of Bushee, Core, Guay and Hamm (2010) when examining the impact of press coverage on information asymmetry around earnings announcements. If more geographically dispersed Internet search decreases information asymmetry, similarly to wider dissemination of earnings news (Bushee, Core, Guay and Hamm 2010), then we should find a reduction in abnormal spreads for higher *HQDisGD*. We estimate the following model,

$$\begin{aligned} AbSpreads\,[0, 1] = {} & a_0 + a_1 R\_absSUE + a_2 R\_HQDisGD + a_3\, chSVI + a_4 RV \\ & + a_5 RECPRC + a_6 absCAR[0,1] + a_7 Turnover + a_8 BM + a_9 Ln(SHR) \\ & + a_{10} Ln(Size) + a_{11}\#\, of\, News + a_{12} IO + a_{13} LnAF + a_{14} MF + \varepsilon, \end{aligned} \tag{3}$$

where *AbSpreads[0,1]* is defined in Section 3.3; *R_absSUE*, and *R_HQDisGD* are defined above; following Drake, Roulstone and Thornock (2012) *chSVI* is the natural logarithm of the firm's weekly search volume index (SVI) minus the median value of the firm's SVI over the previous ten weeks; *RV* is the standard deviation of daily returns from the prior quarter; *RECPRC* is the reciprocal of stock price; *absCAR[0,1]* is the absolute value of cumulative abnormal returns over the two-day earnings announcement window; *Turnover* is the quarterly average monthly trading volume for the stock scaled by total shares outstanding; and *BM*, *Ln(SHR)*, *Ln(Size)*, *# of News*, *IO*, *LnAF* are defined as described above (computed on a quarterly basis).

Table 4 reports results. Our prediction focuses on the coefficient of *R_HQDisGD*. As expected, we find a significantly negative coefficient on *R_HQDisGD*, significant at the 1% level in all four models. The standard deviation of *AbSpreads[0,1]* is 0.386 (Table 2). Thus moving from the bottom to top decile of search dispersion is associated with a drop of 2.85% (column 3) to 3.89% (column 1) of a standard deviation in abnormal spreads during the earnings announcement window. By comparison, Bushee, Core, Guay and Hamm (2010, Tables 1 and 3) find that a one standard deviation change in press coverage is associated with a change in abnormal spreads of 6.5% of a standard deviation. Thus, the magnitude of the effect for geographic dispersion of search is slightly smaller than, but on the same order of magnitude as, the effect for press coverage. These results suggest that more dispersed search reduces information asymmetry (lower bid-ask spreads), incremental to factors examined in prior literature. In contrast, higher abnormal search volume, *chSVI*, are associated with significantly *higher* bid-ask spreads, and higher absolute earnings surprises are associated with insignificantly or marginally significantly higher bid-ask spreads as prior literature documented.

[Insert Table 4]

Next we examine the relation between the geography of search and abnormal trading volume (*AbVol[0,1])*. We employ a specification similar to the one we use to analyze *AbSpreads[0,1]*. We regress *AbVol[0,1]* on the quarterly decile ranks of absolute value of earnings surprise (*R_absSUE*), the quarterly decile ranks of lagged *HQDisGD* (*R_HQDisGD*), and a set of control variables. We estimate the following model,

$$\begin{aligned} AbVol\,[0, 1] = {} & \sigma_0 + \sigma_1 R\_absSUE + \sigma_2 R\_HQDisGD + \sigma_3\, chSVI + \sigma_4 BM \\ & + \sigma_5 Ln(Size) + \sigma_6 \#\ of\ News + \sigma_7 IO + \sigma_8 LnAF + \sigma_9 MF \\ & + \sigma_{10} R\_EA + \sigma_{11} EV + \sigma_{12} EP + \sigma_{13} MKVOL[0,1] + \varepsilon, \end{aligned} \tag{4}$$

where *AbVol[0,1]* is defined in Section 3.3; *R_absSUE, R_HQDisGD, chSVI, BM, Ln(Size), # of News, IO, LnAF,* and *MF* are defined as described above; *R_EA* is the normalized quarterly decile rank of the number of earnings announcement of other firms on the same day (Hirshleifer, Lim and Teoh 2009); *EV (EP)* is quarterly earnings volatility (persistence) over the past four years (Drake, Roulstone and Thornock 2012); and *MKVOL[0,1]* is the market-wide average daily trading volume over the earnings announcement window.

We predict a positive value for $\hat{\sigma}_2$, consistent with *R_HQDisGD* increasing abnormal trading volume around earnings announcements. Results are displayed in Table 5. The coefficient on *R_HQDisGD* is positive and significant, with p-values of 0.004, 0.014, 0.042 and 0.006 in columns 1-4, respectively. To put the coefficient estimates into perspective, recall that the standard deviation of *AbVol[0,1]* is 0.590 (Table 2). Focusing on column 1, the coefficient on *R_HQDisGD* is 0.005. Increasing from the lowest to the highest decile of *HQDisGD* would increase abnormal volume by 0.85% of a standard deviation. Across columns 1-4, the magnitude of the effect for *R_HQDisGD* is similar to the effect of a one standard deviation change in the abnormal search level for the firm, *chSVI*, and is roughly 2% to 5% of the effect of moving from the lowest to highest deciles of absolute earnings surprise, *R_absSUE,* which is one of the strongest drivers of abnormal trading volume around earnings announcements (Beaver 1968; Bamber 1987; Bamber, Barron and Stevens 2011). Overall, the results indicate that the geographic distribution of search is significantly related to abnormal volume.

[Insert Table 5]

Overall, the results displayed in Tables 4 and 5 provide evidence that broader geographic dispersion of search is associated with lower information asymmetry, as captured by abnormal bid-ask spreads and trading volume, around earnings announcements. These results are consistent with the predictions of network theory that having more dispersed search will lead to broader information diffusion, similarly to broader press coverage.

4.2.2 The Incorporation of Information into Price: Returns around Earnings Announcements

We predict that the market reaction to earnings announcements will be more complete if geographic dispersion of search is broader. *Network theory* suggests that more dispersed search leads to the information reaching more investors, while *psychological theory* suggests that more dispersed search leads to the information reaching investors who will respond more rationally to it. Both lead to the prediction of a stronger and faster returns response to earnings information for firms with more dispersed search. In this subsection we examine the market reaction to earnings announcements, to test these predictions. Specifically, we are interested in how the earnings response coefficient (ERC), as captured by the relation between *CAR[0,1]* and earnings news, and post-earnings-announcement drift (PEAD), as captured by the relation between *CAR[2,61]* and earnings news, vary with *HQDisGD*. We estimate the following regression to examine the initial earnings response,

$$CAR\ [0,1] = \sigma_0 + \sigma_1 R\_SUE + \sigma_2 R\_HQDisGD + \sigma_3\ R\_SUE*R\_HQDisGD + \sigma_4\ Controls + \sigma_5 R\_SUE*Controls + \varepsilon, \quad (5)$$

where *CAR[0,1]* is defined in Section 3.3; *R_SUE* is the quarterly decile rank of standardized unexpected earnings; *Controls* is a set of variables including *chSVI, BM, Ln(Size), R_EA,# of News, IO, LnAF, Turnover, MF, EV* ,and *EP,* as defined above. We predict that *CAR[0,1]* is more strongly related to earnings news for firms with higher *HQDisGD*, suggesting a positive value for $\hat{\sigma}_3$.

Table 6 presents the results from estimating equation (5). First, consistent with the ERC literature, announcement-window returns, *CAR [0,1],* are positively associated with the earnings surprise, *R_SUE*. In addition, the result for control variables and their interactions with *R_SUE* (unreported for brevity) are generally consistent with prior literature. The estimated coefficient $\hat{\sigma}_3$ on the interaction term (*R_SUE*R_HQDisGD*) is significantly positive, with p-values of 0.004, <0.001, 0.018 and <0.001 in columns 1-4, respectively, using two-tailed tests. Comparing the coefficients on *R_SUE* and *R_SUE*R_HQDisGD*, the earnings response coefficient increases by 8-9% (4-6%) for random-walk-based (analyst-based) earnings surprises, when *HQDisGD* increases from the bottom to top decile.[9]

[Insert Table 7]

We further test the relation between *HQDisGD* and PEAD. If investors around the country are searching for information about the firm, resulting in a stronger ERC as suggested by the results in Table 6, then we predict that PEAD will be weaker. Thus we expect $\hat{c}_3$ to be negative when estimating the following model,

$$CAR[2,61] = c_0 + c_1 R\_SUE + c_2 R\_HQDisGD + c_3\, R\_SUE * R\_HQDisGD \\ + c_4 Controls + c_5 R\_SUE * Controls + \varepsilon, \quad (6)$$

[9] Focusing on the sum of coefficients on *R_SUE, R_HQDisGD,* and *R_SUE*R_HQDisGD*, we find that more geographically dispersed search is associated with 11-15% more negative returns for the most negative earnings surprises (*R_SUE*=-0.5), and 26-38% more positive returns for the most positive surprises (*R_SUE*=0.5). That is, it is associated with lower drift at both ends of the earnings surprise spectrum. Summary statistics for each decile, which do not control for other factors, show an even stronger association.

where *CAR[2,61]* is defined in Section 3.3.; *R_HQDisGD* and *R_SUE* are defined above; and *Controls* is the same set of control variables as in equation (5).

Table 7 reports the results of estimating equation (6). Consistent with the post-earnings-announcement drift literature, we find a positive coefficient on *R_SUE*, implying that firms with more positive earnings surprises generate more positive abnormal returns in the three months after the earnings announcement. The result of control variables and their interactions with *R_SUE* (unreported for brevity) are generally consistent with prior literature. For example, firms followed by more analysts generate lower drift.

As expected, the coefficient on the interaction term *R_SUE*R_HQDisGD* is significantly negative using both random walk earnings surprise (columns 1-2) and analyst forecast earnings surprise (columns 3-4), implying that firms with higher geographic dispersion of search have lower post-earnings-announcement drift. The effect is statistically significant, with p-values of 0.011, 0.046, 0.017 and 0.014 in columns 1-4, respectively, using two-tailed tests. The drop in drift is large, with a drop of 23% (column 1) to 31% (column 4) for an increase from the lowest to the highest decile of *HQDisGD*. This result shows that subsequent post-earnings-announcement drift is significantly weaker when Google search for the firms' information is more geographically dispersed.[10] It is interesting to note that there is no significant drop in drift associated with higher levels of abnormal search or greater news coverage. While news coverage and search levels are associated with

[10] Focusing on the sum of coefficients on *R_SUE*, *R_HQDisGD*, and *R_SUE*R_HQDisGD*, we find that more geographically dispersed search is associated with 13-20% less negative returns for the most negative earnings surprises (*R_SUE*=-0.5), and 28-34% less positive returns for the most positive surprises (*R_SUE*=0.5). That is, it is associated with lower drift at both ends of the earnings surprise spectrum. Summary statistics, which do not control for other factors, show the complete elimination of drift in many cases. For example, firms in the lowest earnings surprise decile earn negative *CAR[2,61]* on average for *HQDisGD* in deciles 1-6, but positive *CAR[2,61]* for *HQDisGD* in deciles 7-10.

higher abnormal trading volume, and abnormal search is associated with higher earnings response coefficients, neither is associated with lower drift.

The results reported in this section are consistent with more geographically dispersed investor interest enhancing the initial market reactions to firms' earnings news, both in terms of lowering information asymmetry and improving the incorporation of information into price, and decreasing subsequent post-earnings-announcement drift, supporting H2 and H3.

[Insert Table 8]

4.3 Additional Analyses

In this section, we report several additional analyses, including an examination of changes in dispersion, comparison of propensity-score matched firms, path analysis to examine alternate mechanisms for the relations we document, and additional sensitivity analyses.

4.3.1 Changes Analysis

If the geographic dispersion of search for a firm is highly persistent, then reverse causality is an important concern, despite our use of lagged dispersion. In this section we examine changes in annual geographic dispersion of search. We examine whether, if search for a firm's ticker becomes more dispersed, earnings-window liquidity, bid-ask spreads, and earnings response coefficients (ERCs) subsequently improve, and post-earnings-announcement drift (PEAD) drops. We sort firms into terciles in each year *t* based on $Ln(HQDisGD_t)–Ln(HQDisGD_{t-1})$. The top (H) tercile experiences a statistically significant average increase in *Ln(HQDisGD)* of 0.24, 31% of a standard deviation, while the bottom (L) tercile experiences a significant drop in dispersion of 0.24, 32% of a standard deviation. For each tercile, we examine abnormal trading volume, abnormal bid-ask spreads, and returns around and after earnings announcements, for the year before the change in search dispersion (Yr_{t-1}) and the year after the change (Yr_{t+1}).

Focusing first on abnormal bid-ask spreads, we estimate equation (3) without *R_HQDisGD* for each tercile, before and after the change in dispersion. We focus on the intercept in the regression, which captures average abnormal bid-ask spreads around earnings announcements for the given group of firm-years, after controlling for other determinants. Table 8, columns 1-3, present the results for the top (H) and bottom (L) terciles. Prior to the change in *Ln(HQDisGD)*, during Yr_{t-1}, there is no significant difference between the announcement-window abnormal bid-ask spreads for the H and L terciles, controlling for other factors. The focus of our analysis is on the third column, the differences from before to after the change, and the difference in difference. We find that firms with the highest increases in geographic dispersion of search experience a statistically and economically significant 86% (69%) drop in abnormal spreads when measuring the earnings surprise using the random-walk (analyst-based) model, from the statistics reported in Panel A (B). Firms in the bottom category, with significant decreases in dispersion of search, experience a smaller drop in abnormal spreads (Panel A), or no drop (Panel B). The difference between the two groups is significant with $p=0.014$ ($p=0.002$), using a two-tailed test, in Panel A (B).

We use a similar method for abnormal trading volume, estimating a regression similar to equation (4), but with the *R_HQDisGD* term removed. Table 8 shows an increase in abnormal trading volume of almost double the original volume for firms with the largest increases in search dispersion, with $p<0.001$ in both panels. There is no significant change in abnormal trading volume for firms in the L category, and the difference in difference is significant at the 1% level in both panels.

To examine the change in the ERC, we estimate equation (5), removing the *R_HQDisGD* and *R_SUE*R_HQDisGD* terms. We focus on the change in the coefficients of *R_SUE*, capturing the ERC, from before to after the change in dispersion. The ERCs are significantly positive for all groups, but increase significantly for the high (H) change in dispersion group (by 62-66% of the pre-change value, $p \leq 0.005$) and remain unchanged for the low (L) group. The differences in difference are significant with $p=0.003$ in each panel.

We use a similar method to examine PEAD. PEAD is positive and significant for each group, but drops by 61-62%, $p \leq 0.006$, for the H group, and changes insignificantly for the L group, with a significant difference between the two groups.

Overall we find an economically and statistically significant decrease in information asymmetry, increase in liquidity, increase in ERC, and decrease in PEAD for firms with the highest increases in geographic dispersion of search, both in absolute terms, and when compared to firms with the largest drops in geographic dispersion. Thus, reverse causality is unlikely to drive our results, as are persistent firm characteristics.

4.3.3 Propensity Score Matched-Pair Analysis

Several of our research design decisions, such as using the firm as its own control, including control variables for dimensions of firm visibility which might be related to our variables of interest, and conducting a changes analysis, help to control for the impacts of cross-sectional variation in visibility on the market reaction around earnings.[11] However it is

[11] As an additional method to address this issue, we use the residual (*Residual*) from equation (2), which regresses our geographic dispersion of search measure on variables capturing the visibility of the firm. *Residual* captures the variation in *HQDisGD* which is orthogonal to the collection of visibility measures included in equation (2). We use the rank of *Residual* in these tests, just as we use the rank of *HQDisGD* in our primary tests. The results for *AbSpreads[0,1]* are with statistically significant negative coefficients on *R_Residual* and of a similar magnitude to the primary results using the random walk based earnings surprise measure and roughly half the magnitude using the analyst forecast based measure. The results for abnormal

possible that visibility affects the market reaction around earnings in ways that our linear models do not sufficiently control for. In order to address this issue, we use a propensity score matched-pair research design to form firm-year matched pairs that are similar along the set of firm characteristics included in equation (2) (the "covariates") – the characteristics that we expect to be related to the geographic dispersion of investor interest but which may also affect the market reaction to earnings news – but most dissimilar in terms of geographic dispersion of Google search (*HQDisGD*).[12] Recall from Section 4.1 that there is significant variation in *HQDisGD* even after controlling (linearly) for the large set of visibility-related variables we include in equation (2). After matching on these firm characteristics, any difference in information asymmetries and return reactions to earnings news can be more appropriately attributed to differences in the level of geographic dispersion rather than to differences in the other variables.

The matched sample is constructed using a nonbipartite matching algorithm suggested by Derigs (1988) and Lu, Greevy, Xu and Beck (2011) (see, e.g., Armstrong, Blouin and Larcker 2012; Armstrong, Jagolinzer and Larcker 2009). The algorithm creates optimally matched pairs that minimize the average distance between pairs along the set of covariates on which we match. We match within year, without replacement. We examine covariate balance between our matched pairs and find only one covariate, out of twelve, which is significantly different between the two groups at the mean, median, and distribution:

---

trading volume are with coefficient estimates roughly twice the magnitude of our primary results, and statistically significant at the 1% level. The results also indicate a statistically significant 11-17% increase in earnings response coefficient during earnings announcement window, and a statistically significant 28-34% drop in post-earnings-announcement drift. Thus, our results are robust.

[12] Note that we do not include all of these control variables in each of equations (3), (4), (5), and (6), because we do not expect all of them to be related to the different dependent variables in these equations. However we include the full set of variables in our propensity score matching to obtain a single set of matched pairs to analyze.

Ln(SHR).[13] Our results are robust to excluding firm pairs with the largest differences in *Ln(SHR)* to obtain covariate balance,[14] however we tabulate results for the full sample of matched pairs.[15] Similar to the change analysis in Section 4.3.1, we compare the differences in the intercepts of equation (3) for *AbSpreads[0,1]* and equation (4) for *AbVol[0,1]* between treatment firms with relatively high *HQDisGD* and control firms with relatively low *HQDisGD*. We also compare the differences of the coefficients on earnings surprise in equations (5) and (6) for the earnings announcement window and post announcement window. We estimate the equations excluding the control variables which are also used in the propensity-score matching. We further estimate the equations either excluding the additional control variables, to obtain average effects for each of the matched and control samples, or including the additional control variables, to estimate differences between the two groups while controlling for other factors.

[Insert Table 9]

[13] We examine the covariate balance between the treatment (high *HQDisGD*) and control (low *HQDisGD*) samples by comparing the mean and median values of each covariate between the treatment and control samples and testing for differences using a t-test, Wilcoxon Z-test, and Kolmogorov-Smirnov test. We find no significant differences in *Ln(SVI), Urban, Ln(Size), Adv Exp, Ln(EMP), Ln(AF), SP500, Retail,* and *# of News* for the subsample with that data available. *IO* differs using the Kolmogorov-Smirnov test, however the mean values are identical for the two samples and medians differ by only 5.8%. The distrubition and median of *BM* show statistically significant differences. However the treatment sample has *higher* mean and median book to market than the control sample, counter to intuition that "glamour" or growth stocks will have higher visibility and thus higher geographic dispersion. Finally, we find that high *HQDisGD* firms have 26% (16%) higher mean (median) *Ln(SHR)* than control firms, with statistically significant differences in the mean, median, and distribution.

[14] We drop the 20% of matched firm pairs with the poorest match along the dimension of *Ln(SHR)*, resulting in statistically insignificant differences in *Ln(SHR)* between the two groups. Results are robust.

[15] The value for *SHR* that is reported in Compustat is based on the number of distinct shareholders of record. In many cases, the shareholder of record is a brokerage house, as the shares are held in "street name," rather than under the investor's name. Thus many factors other than the actual number of shareholders can affect this variable. Because of this, and the strong match between treatment and control firms along all other dimensions, we choose to include the full set of matched pairs in our main tests and exclude the *Ln(SHR)* unbalanced pairs as a robustness test, rather than the reverse.

The results are reported in Table 9. Table 9 Panel A shows that the treatment firms, on average, have 39-57% lower abnormal bid-ask spreads than the control firms. The difference in abnormal trading volumes is statistically significant in the expected direction in all four models. The magnitude of the difference is small when examining analyst forecast based earnings surprises without controls, at 2.5%, however it is large in the other three models, ranging from 28% to 65%. Panel B indicates that the earnings response coefficient during the announcement window (the relation between *CAR(0,1)* and the earnings surprise) is significantly higher (29-31% higher without controls and 112% higher with controls) for the treatment firms than for the control firms. Finally, post-earnings-announcement drift (the relation between *CAR(2,61)* and the earnings surprise) is significantly lower for the treatment firms than for the control firms (24% drop using random walk based earnings surprise with controls, and 78-88% drop in the other three models). Thus, the results from propensity-score matched pairs are consistent with our primary results, suggesting that other dimensions of visibility, including non-linearity in the relationship between visibility and the market reaction around earnings announcements, are not driving the results.

4.3.3 Path Analysis: Testing Alternate Hypothesized Mechanisms Linking Search Dispersion with Market Reactions to Earnings

Sections 4.2, 4.3.1, and 4.3.2 all show a statistically and economically significant relation between geographic dispersion of search and ERCs and PEAD. Both *network theory* and *psychological theory* predict this relation, however the mechanisms for the two differ. *Network theory* predicts that when information reaches more networks it will reach more investors. This will lead to more investor trading, which in turn improves the market reaction to earnings announcements. Thus, if network effects drive this relation, then trading volume around the earnings announcement window should be part of the mechanism linking dispersion to returns. *Psychological theory* predicts that when information reaches distant investors, those investors will incorporate the information better, for example by making the appropriate comparisons with other firms and by thinking more long-term, and thus make "smarter" (i.e. more rational) trades. Investors making smarter trades will improve the overall market reaction. Thus, if investor psychology drives the relation, then "smarter trading" should be part of the mechanism, as distant investors respond more appropriately. In this section we conduct a path analysis to test the importance of each of these mechanisms. We use path analysis to measure the direct effect of geographic dispersion on returns and the indirect effects through trading volume and investor education (using the educational attainment in search states as a proxy for the likelihood of "smarter trading").

Path analysis is a method for testing the importance of different mediating variables. We examine the importance of trading volume during the earnings announcement window *(Vol[0,1])* as a mechanism, to test network theory. To test psychological theory, we use the educational attainment of individuals in the states searching for a firm (*College_Ed*) as a proxy for investor sophistication.[16] We make the simplifying assumption that paths flow in one direction, as in Landsman, Maydew and Thornock (2012).

Path analysis separates the correlation between two variables into a direct path and an

[16] Ideally we would like a measure of the sophistication of investors who use Google search to obtain information for each firm. We want a measure of investor sophistication that is not based on stock price behavior, since stock price behavior may be driven by other factors, such as short-sale costs and constraints, which would also impact the market response to earnings. We obtain data for the percentage of individuals in each state with college degrees, as of 2004 and 2009, from the U.S. Census Bureau. Because educational attainment tends to be persistent, we use the average of the 2004 and 2009 rates to calculate educational attainment for each state during our sample period. For each firm, we then calculate a weighted average educational attainment. Specifically, we weight the state attainment rates by the firm's search volume index for each state. It is possible that it is the most educated individuals within a state that use Google to search for distant firms, which would decrease our power to identify education-related effects. However our measure serves as a proxy for the educational attainment of those searching for a firm's ticker on Google.

indirect path through mediating variables (Baron and Kenny 1986; MacKinnon, Fairchild, and Fritz 2007). A direct path contains only one path coefficient while an indirect path contains a path coefficient between the source variable and the mediating variables and between the mediating variables and the dependent variable. In our setting, the indirect paths are between the interaction term of *R_SUE*R_HQDisGD* and *Vol[0,1]* or *College_Ed* and between *Vol[0,1]* or *College_Ed* and the return reactions to earnings news (*CAR[0,1]* or *CAR[2,61]*). We estimate the following structural equation model,

$$CAR\,[0,1]\ or\ CAR\,[2,61] = a_0 + a_1{*}R\_SUE + a_2{*}R\_SUE{*}R\_HQDisGD + a_3{*}R\_SUE{*}Vol[0,1] + a_4{*}R\_SUE{*}College\_Ed + a_5{*}Vol[0,1] + a_6{*}College\_Ed + a_x{*}Controls + a_y{*}Controls{*}R\_SUE + \varepsilon, \quad (7)$$

$$Vol\,[0,1]\ \text{or}\ College\_Ed = b_0 + b_1{*}R\_HQDisGD + b_X{*}Controls + \varepsilon, \quad (8)$$

where all variables are standardized to mean zero and a standard deviation of one to allow comparisons between indirect and direct path coefficients.[17] The estimated value $\hat{a}_2$ measures the magnitude of the direct path effect from the geographic dispersion of search to the returns associated with a given earnings surprise. The product of $\hat{b}_2$ estimated with *Vol[0,1]* (*College_Ed*) and $\hat{a}_3$ ($\hat{a}_4$) measures the magnitude of the indirect path from geographic dispersion to returns through trading volume (investor education). Table 10 reports the results for the path coefficients, for both the immediate return response to earnings *(CAR[0,1])* and post-earnings-announcement returns *(CAR[2,61])*. The results in all four columns show significant direct and indirect effects. The direct path between *HQDisGD* and returns accounts for 58-66% of the total effect. The trading volume path accounts for 20-26% of the effect. Broader geographic dispersion of search is related to higher trading volume,

[17] In equation (7), the control variables are *chSVI*, *BM*, *Ln(Size)*, *R_EA*, *LnAF*, *MF*, *EV*, *EP*, and *IO*. In equation (8), the control variables are *chSVI, BM, Ln(Size), LnAF*, and *IO*. See Appendix A for variable definitions.

which in turn are related to higher ERCs and lower drift. Broader geographic dispersion of search is also related to higher search-state educational attainment, which in turn affects the market response to earnings. Investor education accounts for 12-22% of the effects. Thus, overall, we find significant evidence in support of both the *network* and *psychological theories* for the importance of the geographic dispersion of search. The remaining direct effect that we estimate may be related to either of these theories as well, if our proxies fail to capture the full effects of increased investor awareness of the information (*network theory*) or smarter distant-investor trading (*psychological theory*).[18]

[Insert Table 10]

4.3.4 Sensitivity Analyses

Our primary measure has several advantages, such as cancelling out the Google Trends scaling factor, and measuring distance similarly to the prior local bias literature. However to ensure that our results are robust to alternate measures of the geographic dispersion of search, for example capturing breadth rather than distance, we conduct the following specific tests. First, we use the sum of state-specific SVI, which is a simple non-distance-based measure of the breadth of search. Second, instead of distance, we weight each state-specific SVI by the area of the state. This gives us a measure of the geographic breadth of search adjusting for the fact that some states (e.g., Texas, California) are significantly larger than others. Third, we use the firm's top search state in a given year as the "center" for the geographic-distance calculation, rather than using the firm's headquarter location. We

[18] In untabulated analysis, we replicate these tests using an alternate measure of "investor education." We use the percent of college-educated individuals in each state who have their degree in business since those individuals are arguably the most likely to react appropriately to the earnings information. Using this measure we find a slightly higher total impact for the investor education path. The direct path accounts for 41-61% of the total effect, the trading volume path accounts for 19-24%, similarly to before, and the investor education path accounts for 15-38% of the total effect.

also replace the rank decile variable, *R_HQDisGD*, with the raw value of *Ln(HQDisGD)* and, alternatively, replace the lagged *R_HQDisGD* match with contemporaneous *R_HQDisGD*. Results for *AbVol[0,1]*, *CAR[0,1]* and *CAR[2,61]* are robust in all five variations. Results for *AbSpreads[0,1]* marginally lose significance when we use state area size as the weight and distance based on the geographic center of firm's top search state instead of firm's headquarter location, with 2-tailed p-values of 0.106 and 0.112, respectively. Overall, the results remain consistent.

Regarding other variables, we use an alternate method to compute abnormal spreads using daily high and low prices, as in Corwin and Schultz (2012). Replicating Table 4 with this variable, we find that the coefficient on *R_HQDisGD* is significantly negative at the 5% level or better in all four models. We also verify that results are robust to smaller changes in variable definitions, such as using the average of the highest and lowest trading price during the day to calculate $\$vol_t$ rather than closing price, and using a 60 trading day window to calculate *AbVol* and *AbSpreads* rather than 40 trading days.

To examine the evolution of post-earnings-announcement returns, we conduct tests for one month, six month and one year windows. The dispersion-related drop in drift, captured by the coefficient estimate for *R_SUE*R_Ln(HQDisGD)*, is largest for the one month window, and is less negative and statistically insignificant for the six-month and one-year windows. Together, the results suggest that geographically dispersed search increases the initial response to earnings announcements, decreasing post-earnings-announcement drift for up to three months after the earnings announcement. However after that point, return differences related to the dispersion of search diminish and are subsumed by other factors.

Finally, we re-estimate our tests excluding certain subsets of the data. First, we exclude firms located in Washington, as Microsoft's Bing search engine is likely to be more popular in Washington State. This would lead to understating the amount of search for Washington firms in their home state, since we use Google search data. Second, we exclude financial and technology firms. Financial firms tend to be geographically concentrated around New York City, while technology firms tend to be geographically concentrated in California. Both may have unique search characteristics. Third, we exclude firms headquartered in California, Texas and New York, as one-third of our sample firms are located in these three large states. Our main results and inferences remain unchanged.

5. Conclusion

We examine whether geographic proximity affects investors' search behavior for firms during the Internet era. We find strong evidence that investors search more heavily for companies headquartered near them than for distant companies. This local bias of investor search is of similar magnitude to local bias of ownership during the early 1990's, and is strongest for firms that are less visible to investors (e.g., small firms, rural firms, non-retail firms). We also find that the geographic dispersion of Internet search impacts the market response around earnings announcements. Higher geographic dispersion of search is related to lower abnormal bid-ask spreads and higher abnormal trading volume around earnings announcements, suggesting a relative reduction in information asymmetry for firms with a broader geographic dispersion of interested investors. Firms with more dispersed search also experience stronger return responses to earnings news around the earnings announcement and weaker post-earnings-announcement drift, suggesting that broad geographic interest in a firm accelerates the incorporation of information into prices. Firms with the largest annual increases in search dispersion experience a significant improvement in the market response around their earnings announcements. Finally, a path analysis shows that both network effects, through the mechanism of trading volume, and investor psychology, through the mechanism of investor education, contribute to the return results.

While prior literature has established that location affected investor behavior and the incorporation of information into stock price before the dramatic rise of the Internet, our paper sheds light on whether distance is still important in the Internet era. While the Internet has made it easier for investors to search for information about distant firms, geography still plays an important role in investor interest and impacts the market response to information. Furthermore, our study sheds light on both *network theory* and *psychological theory* for the importance of geography. Our results provide evidence in support of the idea that information is diffused within local networks, and that network effects contribute to the importance of dispersed search. Our findings also indicate that physical distance affects investors' ability to process information, consistent with "psychological distance" leading more local investors to make less sophisticated investing choices.

Practitioners with an interest in increasing firm visibility and improving the market response to firm information, such as executives and investor relations professionals, will benefit from realizing the importance of national firm recognition. The results of our analysis of changes in search dispersion suggest that firms can improve the market reaction to their earnings information by increasing the level of non-local investor interest. Increasing the number of interested investors is important, however increasing the geographic breadth of investors interested in the firm also matters. Researchers interested in the dissemination of information, by investors or information intermediaries (e.g., the press, analysts), will also want to consider the role of geography. Finally, our results reinforce the message of studies such as Kedia and Rajgopal (2011) and DeFond, Francis and Hu (2011) that monitors, such as auditors or the SEC, may want to consider the role of geography in their own monitoring activities. Our results indicate that geography can be of significant importance even in the Internet age, due to the role that geography plays in where individuals direct their attention.

References

Armstrong, C.S., Jagolinzer, A.D., Larcker, D.F., 2009. Chief executive officer equity incentives and accounting irregularities. *Journal of Accounting Research* 48, 225–271.

Armstrong, C., Blouin, J., D. Larcker, 2012. The incentives for tax planning. *Journal of Accounting and Economics* 53(1): 391-411.

Ayers, B., C., O. Z. Li, and P. E. Yeung, "Investor trading and the post-earnings-announcement drift." *The Accounting Review* 86 (2011): 385-416.

Ayers, B. C., S. Ramalingegowda, and P. E. Yeung. "Hometown advantage: The effects of monitoring institution location on financial reporting discretion." *Journal of Accounting and Economics* 52 (2011): 41-61.

Bamber, L.S.. "Unexpected earnings, firm size, and trading volume around quarterly earnings announcements." *The Accounting Review* 62 (1987): 510–532.

Bamber L.S., Barron O., and Stevens D.E. "Trading volume around earnings announcements and other financial reports: theory, research design, empirical evidence, and directions for future research *Contemporary Accounting Research* 28(2011): 431-471.

Barber, B., and T. Odean. "Trading is hazardous to your wealth: The common stock investment performance of individual investors." *Journal of Finance* 55 (2000): 773-806.

Baron, R., and D. Kenny. "The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations." *Journal of Personality and Social Psychology* 51 (1986): 1173–1182.

Bartov, E., I. Krinsky, and S. Radhakrishnan. "Investor sophistication and patterns in stock returns after earnings announcements." *The Accounting Review* 75 (2000): 43–63.

Battalio, R., H., and R. R. Mendenhall. "Earnings expectations and investor clienteles." *Journal of Financial Economics* 77 (2005): 289-319.

Bhattacharya, N. "Investors' trade size and trading responses around earnings announcements: An empirical investigation." *The Accounting Review* 76: 221-244.

Beaver, W. 1968. ''The information content of annual earnings announcements.'' *Journal of Accounting Research* 6 (Supplement): 67–92.

Bernard, V., and J. Thomas. "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" *Journal of Accounting Research* 27 (1989): 1-36.

Bernard, V., and J. Thomas. “Evidence that stock prices do not fully reflect the implications of current earnings for future earnings.” *Journal of Accounting and Economics* 13 (1990): 305–340.

Brown, T., and B. Shetty. “A Turn of the Page for Newsweek. After 80 years in print, the newsmagazine adopts an all-digital format.” *Newsweek/The Daily Beast*, October 18, 2012, http://www.thedailybeast.com/articles/2012/10/18/a-turn-of-the-page-for newsweek.html.

Bushee, B., J. Core, W. Guay, and S. Hamm. “The role of the business press as an information intermediary.” *Journal of Accounting Research* 48 (2010): 1-19.

Bushee, B., and G. S. Miller. “Investor relations, firm visibility, and investor following.” *The Accounting Review* 87 (2012): 867-897.

Collins, D., and S. Kothari. “An analysis of intertemporal and cross-sectional determinants of earnings response coefficients.” *Journal of Accounting and Economics* (1989): 143-181.

Corwin, S. A., and P. Schultz. “A simple way to estimate bid-ask spreads from daily high and low prices.” *The Journal of Finance* 67 (2012): 719-759.

Coval, J. D., and T. J. Moskowitz. “Home bias at home: local equity preference in domestic portfolios.” *The Journal of Finance* 54 (1999): 2045-2073.

Coval, J. D., and T. J. Moskowitz. “The geography of investment: Informed trading and asset prices.” *Journal of Political Economy* 109 (2001): 811-841.

Da, Z., J. Engelberg, and P. Gao. “In search of attention.” *Journal of Finance* 66 (2011): 1461-1499.

DeFond, M. L., J. R. Francis, and X. Hu. “The geography of SEC enforcement and auditor reporting for financially distressed clients.” Working Paper, University of Southern California, University of Missouri at Columbia, and University of Oregon, 2011.

Derigs, U.,1988. Solving non-bipartite matching problems via shortest path techniques. *Annals of Operations Research* 13,225–261.

Drake, M. S., N. M. Guest, and B. J. Twedt. “The media and mispricing: The role of the business press in the pricing of accounting information.” *The Accounting Review* forthcoming (2014).

Drake, M. S., D. T. Roulstone, and J. R. Thornock. “Investor information demand: Evidence from Google searches around earnings announcements.” *Journal of Accounting Research* 50 (2012): 1001-1040.

Engelberg, J. E., and C. A. Parsons. “The causal impact of media in financial markets.” *The*

Journal of Finance 66 (2011): 67-97.

Fama, E. F., and K. R. French. “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics* 33 (1993): 3-56.

Fama, E. F., and K. R. French. “Industry costs of equity.” *Journal of Financial Economics* 43 (1997): 153–193.

Gow, I., G., Ormazabal, and D. Taylor. “Correcting for cross-sectional and time-series dependence in accounting research.” *The Accounting Review* 85 (2010): 483-512.

Grullon, G., G. Kanatas, and J.P. Weston. “Advertising, breadth of ownership, and liquidity.” *Review of Financial Studies*. 17 (2004): 439-461.

Hirano, K., and G.W. Imbens. “The Propensity Score with Continuous Treatments.” in *Applied Bayesian Modeling and Causal Inference From Incomplete Data Perspectives*, edited by A. Gelman, and X. L. Meng. England: West Sussex, 2004: 73–84.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. “Driven to distraction: extraneous events and underreaction to earnings news.” *The Journal of Finance* 64 (2009): 2289-2325.

Hirshleifer, D., J. N. Myers, L. A. Myers, and S. H. Teoh. “Do individual investors drive post-earnings-announcement drift? Direct evidence from personal trades.” *The Accounting Review* 83 (2008): 1521-1550.

Hong, H., J. D. Kubik, and J. C. Stein. “Social interaction and stock-market participation.” *The Journal of Finance* 59 (2004): 137-163.

Hong, H., J. D. Kubik, and J. C. Stein. “Thy neighbor’s portfolio: Word-of-mouth effects in the holdings and trades of money managers.” *The Journal of Finance* 60 (2005): 2801-2824.

Hong, H., J. D. Kubik, and J. C. Stein. “The only game in town: Stock-price consequences of local bias.” *Journal of Financial Economics* 90 (2008): 20-37.

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

Ke, B., and K. Petroni. “How informed are actively trading institutional investors? Evidence from their trading before a break in a string of consecutive earnings increases.” *Journal of Accounting Research* 42 (2004): 895-927.

Kedia, S., and S. Rajgopal. “Do the SEC’s enforcement preferences affect corporate misconduct?” *Journal of Accounting and Economics* 51 (2011): 259-278.

Lakonishok, J., A. Shleifer, and R. W. Vishny. “Contrarian investment, extrapolation, and risk.” *Journal of Finance* 49 (1994): 1541-1578.

Landsman, W., E.L. Maydew, and J.R. Thornock. “The information content of annual earnings announcements and mandatory adoption of IFRS.” *Journal of Accounting and Economics* 53 (2013): 34-54.

Livnat, J., and R. R. Mendenhall. “Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts.” *Journal of Accounting Research* 44 (2006): 177–205.

Lou, D. “Attracting investor attention through advertising.” *Review of Financial Studies.* Forthcoming.

Loughran, T., and P. Schultz. “Liquidity: Urban versus rural firms.” *Journal of Financial Economics* 78 (2005): 341-374.

Lu B, Greevy R, Xu X, Beck C. Optimal Nonbipartite Matching and its Statistical Applications. *The American Statistician*. Vol. 65, no. 1. : 21-30. 2011.

MacKinnon, D., Fairchild, A., Fritz, M. “Mediation analysis.” *Annual Review of Psychology* 58 (2007): 593–614.

Malloy, C. "The geography of equity analysis." *The Journal of Finance* 60 (2005): 719-755.

Merton, R. “A simple model of capital market equilibrium with incomplete information.” *The Journal of Finance* 42 (1987): 483-510.

Miller, G., and D. M. Shanthikumar. “The press and local information advantage.” Working Paper, University of Michigan at Ann Arbor and University of California Irvine, 2012.

Petersen, M. “Estimating standard errors in finance panel data sets: Comparing approaches.” *Review of Financial Studies* 22 (2009): 435–480.

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

Shanthikumar, D. “Consecutive earnings surprises: Small and large trader reactions.” *The Accounting Review* 87 (2012): 1709-1736.

Shores, D. “The association between interim information and security returns surrounding earnings announcements.” *Journal of Accounting Research* 28 (1990): 164-181.

Soltes, E. “News dissemination and the impact of the business press.” Working paper, Harvard University, 2010.

Stoll, C. “The Internet? Bah. (computers cannot replace books, teachers or newspapers).” Newsweek 125 (February 27, 1995): 41.

Teoh, S., and T. Wong. “Perceived auditor quality and the earnings response coefficient.” *The Accounting Review* 66 (1993): 346-366.

Trope, Y., and N. Liberman. “Construal-Level Theory of Psychological Distance.” *Psychological Review* 117 (2010): 440-463.

Appendix A. Variable Definitions

Variable	Definition
# of News	The number of articles in the *Wall Street Journal*, the *New York Times*, *USA Today*, and the *Washington Post* that mention the firm on the earnings announcement date from 2004 through 2008. Annual basis for Table 3. Quarterly basis for Table 4 onward. (Soltes 2010)
AbSpreads[0,1]	The difference between average bid-ask spreads over the earnings announcement window [0, 1] and the average bid-ask spreads over days [-41, -11]. Day 0 is the earnings announcement date. Bid-ask spread is the difference between an offer price and a bid price divided by the midpoint of the offer and bid price (and multiplied by 100). (CRSP)
AbVol [0,1] *Vol [0,1]*	The difference of average log dollar volume between the earnings announcement window [0, 1] (*Vol [0,1]*) and the period over days [-41, -11]. Day 0 is defined as the above. Daily dollar trading volume is the product of the daily closing price and the daily number of shares traded. (CRSP)
Adv Exp	Advertising expense scaled by sales. Annual data for Table 3. Quarterly data for Table 4 onward. (Compustat)
BM	Book value of common equity divided by market value of equity (size). Annual basis for Table 3. Quarterly basis for Table 4 onward. (Compustat)
CAR[0,1] *AbsCAR[0,1]*	*CAR[0,1]*: Size- and book-to-market-adjusted cumulative abnormal returns over the earnings announcement window [0, 1]. Day 0 is defined as the above. (CRSP) *AbsCAR[0,1]*: The absolute value of *CAR[0,1]*.
CAR[2,61]	Size- and book-to-market-adjusted cumulative abnormal returns over the post-announcement drift period [2, 61]. Day 0 is defined as above. (CRSP)
College_Ed	The SVI-weighted average of state educational attainment. For each firm-year, it is defined using the firm-year's SVI values, and the average educational attainment for the state using 2004 and 2009 Census data. (Google Trends; 2004 and 2009 Census)
EA	The number of other firms announcing quarterly earnings on the same day as the firm announces earnings. (IBES)
EP	The earnings persistence (B_1) over the past four year. B_1 is from the following regression: $\text{Earnings}_t = B_0 + B_1 {*} \text{Earnings}_{t-1} + e$. (Compustat)
EV	Standard deviation of earnings over the past four years. (Compustat)
HQDisGD *Ln(HQDisGD)*	The sum of state Search Volume Index (SVI) weighted by the distance between firms' headquarter locations and geographic center of the state, excluding Alaska and Hawaii, and divided by the sum of state SVI (Google Trends; 2010 Census; Compustat) *Ln(HQDisGD)*: Log of *HQDisGD*.
IO	Percent of shares owned by institutions. Annual basis for Table 3. Quarterly basis for Table 4 onward. (Thomson Reuters 13F)
LnAF	Log of 1 plus the median number of analysts following the firm. Annual basis for Table 3. Quarterly basis for Table 4 onward. (IBES)
Ln(EMP)	Log of 1 plus the number of employees. (Compustat)

Ln(SHR)	Log of 1 plus the number of shareholders. (Compustat)
MF	1 if managers issue a forecast between the fiscal quarter end date and the earnings announcement date; 0 otherwise. (FirstCall)
MKVol[0,1]	Market-wide average daily trading volume during earnings announcement window. (CRSP)
Retail	1 if firm is in retail, consumer goods or entertainment industry[19]; 0 otherwise. (Compustat)
RECPRC	Reciprocal of stock price on the fiscal quarter end date. (Compustat)
RV	The standard deviation of daily raw return in the prior quarter. (CRSP)
Size *Ln(Size)*	Stock price times number of shares outstanding. Annual basis for Table 3. Quarterly basis for Table 4 onward. (Compustat) *Ln(Size):* Log of size.
SP500	1 if the firm is in the S&P 500 Index; 0 otherwise. (Compustat)
SUE *Abs_SUE*	The difference between actual EPS and benchmark EPS, scaled by stock price at the fiscal quarter end. The benchmark EPS is the median analyst forecast EPS for *AF_SUE* and is last year's same quarter EPS for *RW_SUE*. (IBES) *Abs_SUE*: The absolute values of *SUE*.
SVI *Ln(SVI)* *chSVI*	*SVI*: The natural logarithm of (1+weekly Search Volume Index.) (Google Trends) *Ln(SVI):* Log of SVI. *chSVI*: The natural logarithm of {1+(weekly SVI - median SVI over the previous ten weeks)}
Turnover	Quarterly average of monthly trading volume scaled by the total number of shares outstanding. (CRSP)
Urban	1 if a firm's headquarter is located in one of the ten largest metropolitan areas; 0 otherwise. (2000 and 2010 Census; Compustat)

[19] SIC code: 0920-0999, 2000-2039, 2040-2047, 2050-2059, 2060-2068, 2070-2079, 2086-2087, 2090-2092, 2095-2099, 2300-2392, 2510-2519, 2590-2599, 2840-2844, 3020-3021, 3100-3111, 3130-3131, 3140-3149, 3150-3151, 3160-3161, 3170-3172, 3190-3199, 3229-3231, 3260, 3262-3263, 3630-3639, 3650-3652, 3732, 3750-3751, 3800, 3860-3861, 3870-3873, 3910-3911, 3914-3915, 3930-3931, 3940-3949, 3960-3965, 3991, 3995, 5200, 5210-5231, 5250-5251, 5260-5261, 5270-5271, 5300, 5310-5311, 5320, 5330-5331, 5334, 5340-5349, 5390-5399, 5400, 5410-5412, 5420-5469, 5490-5499, 5500, 5510-5579, 5590-5700, 5710-5722, 5730-5736, 5750-5799, 5900, 5910-5912, 5920-5932, 5940-5990, 5992-5995, 5999,7800-7833, 7900, 7910-7911, 7920-7933, 7940-7949, 7980, 7990-7999 (Fama and French 1997)

Table 1. Sample Selection and Distribution

Panel A. Sample Selection Process

	Firm-quarter	Firm
NYSE, AMEX and NASDAQ observations with state Google Search Volume Index	36,024	1,601
Ambiguous tickers	(6,835)	(319)
Insufficient Compustat and CRSP data	(6,640)	(291)
Price per share <$1	(952)	(46)
Baseline observations (2005-2011)	21,597	945

Panel B. Sample Distribution by Year

Year	# of Observations	Percent
2005	2,632	12.19%
2006	2,842	13.16%
2007	3,036	14.06%
2008	3,094	14.32%
2009	3,202	14.82%
2010	3,424	15.86%
2011	3,367	15.59%
Total	21,597	100.00%

Panel C. Sample Distribution by Headquarter State

State	# of Observations	Percent
California	2,855	13.22%
Texas	2,461	11.40%
New York	1,866	8.64%
Illinois	1,192	5.52%
Ohio	1,160	5.37%
Others < 5%	12,064	55.86%
Total	21,597	100.00%

Panel D. Sample Distribution by Industry

Industry	# of Observations	Percent
Retail	2,808	13.00%
Business Services	2,302	10.66%
Electronic Equipment	1,125	5.21%
Others < 5%	15,362	71.13%
Total	21,597	100.00%

Table 2. Descriptive Statistics

Panel A. Main Variables

	Mean	Std Dev	Q1	Median	Q3
AbSpreads[0,1]	0.023	0.386	-0.042	-0.003	0.044
AbVol[0,1]	0.637	0.590	0.269	0.623	0.992
AF_SUE	0.000	0.019	0.000	0.001	0.002
CAR[0,1]	0.002	0.077	-0.035	0.000	0.040
CAR[2,61]	0.000	0.160	-0.083	0.002	0.083
Ln(HQDIsGD)	2.183	0.639	1.943	2.190	2.483
RW_SUE	0.001	0.054	-0.004	0.002	0.007

Panel B. Control Variables

	Mean	Std Dev	Q1	Median	Q3
# of Analyst Coverage	9.834	10.790	2	6	14
# of Earnings Announcements	193.255	114.351	100	200	278
# of News	0.095	0.382	0	0	4
AbsAF_SUE	0.007	0.026	0.001	0.001	0.004
AbsCAR[0,1]	0.055	0.057	0.016	0.038	0.075
AbsRW_SUE	0.022	0.064	0.002	0.006	0.015
Adv Exp	0.010	0.023	0.000	0.000	0.010
BM	0.611	0.456	0.315	0.506	0.774
chSVI	1.502	1.070	0.693	1.504	2.251
EMP	20.783	48.422	0.960	4.290	15.181
EP	0.607	0.299	-0.082	0.291	0.613
EV	0.081	0.124	0.022	0.043	0.084
IO	0.462	0.365	0	0.553	0.800
MF	0.323	0.468	0	0	1
MkVOL[0,1]	78.5996	57.2642	29.442	119.326	258.505
RECPRC	0.089	0.128	0.024	0.042	0.088
Retail	0.130	0.337	0	0	0
RV	0.028	0.017	0.017	0.024	0.035
SHR	32.026	102.195	0.601	2.963	14.814
Size	6677.137	19683.152	216.294	1013.437	3695.220
State Internet Usage	0.055	0.043	0.020	0.038	0.074
SP500	0.301	0.459	0	0	1
SVI	3.631	0.555	3.368	3.772	4.031
Turnover	0.040	0.037	0.015	0.030	0.053
Urban	0.356	0.479	0	0	1

Table 3. Determinats of the Geographic Dispersion of Google Search Volume

	Full Sample				*Subsample with Media Coverage Data*			
	(1)		**(2)**		**(3)**		**(4)**	
Ln(SVI)	0.069	***	0.072	***	0.058	***	0.062	***
	(0.019)		(0.020)		(0.023)		(0.023)	
# of News					0.531	**	0.539	**
					(0.240)		(0.238)	
Urban	0.044	***	0.049	***	0.048	***	0.053	***
	(0.012)		(0.017)		(0.018)		(0.018)	
Ln(Size)	0.234	***	0.234	***	0.157	**	0.197	***
	(0.063)		(0.061)		(0.065)		(0.064)	
Adv Exp	0.771		0.587		1.072	*	0.762	
	(0.576)		(0.511)		(0.620)		(0.550)	
Ln(EMP)	0.076	**	0.073	**	0.071	**	0.071	**
	(0.037)		(0.031)		(0.032)		(0.032)	
Ln(SHR)	0.062	**	0.061	**	0.061	*	0.065	**
	(0.031)		(0.031)		(0.033)		(0.031)	
Ln(AF)	0.127	***	0.127	***	0.124	***	0.176	***
	(0.045)		(0.047)		(0.042)		(0.042)	
IO	0.827	***	0.833	***	0.825	**	0.836	**
	(0.320)		(0.323)		(0.393)		(0.389)	
SP500	0.091	**	0.099	**	0.108	**	0.102	**
	(0.044)		(0.043)		(0.054)		(0.052)	
Retail	0.071	**			0.097	**		
	(0.035)				(0.042)			
BM	-0.035	**	-0.037	**	-0.025		-0.026	
	(0.018)		(0.018)		(0.024)		(0.024)	
Intercept	2.065	***	2.021	***	2.128	***	2.142	***
	(0.163)		(0.131)		(0.209)		(0.142)	
Year Dummy	Yes		Yes		Yes		Yes	
State Dummy	Yes		Yes		Yes		Yes	
Industry Dummy	No		Yes		No		Yes	
N	5821		5821		3975		3975	
Adjusted R-square	9.74%		11.57%		7.57%		9.96%	

This table presents results of estimating ordinary least squares regressions. The dependent variable is the log of distance weighted-average geographic dispersion of Google search - *Ln(HQDisGD)*. The distance is measured between firms' headquarter location and geographic center of each state. *Ln(HQDisGD)* is measured contemporaneously. The full sample consists of firms listed on the major exchanges from 2005 through 2011. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by year and firm. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 4. Abnormal Bid-Ask Spreads around Earnings Announcements

	Random Walk Based Absolute Earnings Surprise (RW_absSUE)		*Analyst Forecast Based Absolute Earnings Surprise (AF_absSUE)*	
	(1)	*(2)*	*(3)*	*(4)*
R_AbsSUE	0.015	0.026 *	0.013	0.025 *
	(0.016)	(0.013)	(0.013)	(0.014)
R_HQDisGD	-0.015 ***	-0.013 ***	-0.011 **	-0.014 ***
	(0.005)	(0.004)	(0.006)	(0.004)
chSVI	0.106 ***	0.125 ***	0.154 ***	0.152 ***
	(0.030)	(0.038)	(0.040)	(0.039)
# of News		-0.015 **		-0.018 *
		(0.007)		(0.010)
Intercept	0.118 **	0.081 *	0.078 *	0.064 *
	(0.047)	(0.048)	(0.048)	(0.037)
Controls	Yes	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes	Yes
N	21597	13837	17842	10501
Adjusted R^2	3.30%	4.40%	6.34%	6.91%

This table presents the results of estimating ordinary least squares regressions. The dependent variable is abnormal bid-ask spreads over trading days zero and one around the earnings announcement, where day 0 is the earnings announcement date. *HQDisGD* is measured in year_{t-1}. The full sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. The subsample with media coverage is from 2005 through 2008. The left-most columns display results using absolute earnings surprise (absSUE) values based on random walk expectations, while the right-most columns present results for absSUE calculated using analyst-based expectations. The control variables are *chSVI, BM, Ln(Size), # of News, IO, LnAF, MF, RV, Ln(SHR), RECPRC, absCAR [0, 1],* and *Turnover. Ln(Size), BM, IO* and *LnAF* are measured as of the fiscal quarter end. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by calendar quarter and firm. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 5. Abnormal Trading Volume around Earnings Announcements

	Random Walk Based Absolute Earnings Surprise (RW_absSUE)				***Analyst Forecast Based Absolute Earnings Surprise (AF_absSUE)***			
	(1)		***(2)***		***(3)***		***(4)***	
R_absSUE	0.113	***	0.189	***	0.209	***	0.191	***
	(0.031)		(0.038)		(0.027)		(0.038)	
R_HQDisGD	0.005	***	0.006	**	0.004	***	0.006	**
	(0.002)		(0.003)		(0.002)		(0.002)	
chSVI	0.005	***	0.004	***	0.005	***	0.004	***
	(0.001)		(0.001)		(0.001)		(0.001)	
# of News			0.007	***			0.020	***
			(0.026)				(0.025)	
Intercept	1.237	***	1.471	***	1.241	***	1.428	***
	(0.131)		(0.223)		(0.141)		(0.233)	
Controls	Yes		Yes		Yes		Yes	
State Fixed Effect	Yes		Yes		Yes		Yes	
N	21597		13837		17842		10501	
Adjusted R^2	3.44%		5.74%		5.30%		6.34%	

This table presents the results of estimating ordinary least squares regressions. The dependent variable is abnormal trading volume over trading days zero and one around the earnings announcement, where day 0 is the earnings announcement date. *HQDisGD* is measured in year_{t-1}. The full sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. The subsample with media coverage is from 2005 through 2008. The left-most columns display results using absolute earnings surprise (absSUE) values based on random walk expectations, while the right-most columns present results for absSUE calculated using analyst-based expectations. The control variables are *chSVI, BM, Ln(Size), R_EA, # of News, IO, LnAF, MF, EP, EV* and *MKVOL[0,1]. Ln(Size), BM, IO,* and *LnAF* are measured as of the fiscal quarter end. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by calendar quarter and firm. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 6. Earnings Announcement Window Cumulative Abnormal Returns

	Random Walk Based Earnings Surprise (RW_SUE)		*Analyst Forecast Based Earnings Surprise (AF_SUE)*	
	(1)	*(2)*	*(3)*	*(4)*
R_SUE	0.057 ***	0.069 ***	0.082 ***	0.092 ***
	(0.008)	(0.008)	(0.010)	(0.012)
R_HQDisGD	0.007	0.007	0.008	0.008
	(0.008)	(0.011)	(0.007)	(0.008)
*R_SUE*R_HQDisGD*	0.005 ***	0.006 ***	0.003 **	0.005 ***
	(0.002)	(0.002)	(0.001)	(0.001)
chSVI	0.003	0.006	0.005	0.006
	(0.004)	(0.006)	(0.005)	(0.007)
*R_SUE*chSVI*	0.028 **	0.020 *	0.022 **	0.028 **
	(0.013)	(0.011)	(0.011)	(0.014)
# of News		0.003		0.004
		(0.003)		(0.004)
R_SUE# of News*		0.009		0.025
		(0.012)		(0.017)
Intercept	0.005	0.014 ***	0.009	0.010
	(0.005)	(0.005)	(0.005)	(0.007)
Controls	Yes	Yes	Yes	Yes
*R_SUE*Controls*	Yes	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes	Yes
N	21597	13837	17842	10501
Adjusted R^2	3.91%	4.54%	7.56%	8.43%

This table presents the results of estimating ordinary least squares regressions. The dependent variable is cumulative abnormal return over trading days zero and one around the earnings announcement, where day 0 is the earnings announcement date. *HQDisGD* is measured in $year_{t-1}$. The full sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. The subsample with media coverage is from 2005 through 2008. The left-most columns display results using earnings surprise (SUE) values based on random walk expectations, while the right-most columns present results for SUE calculated using analyst-based expectations. The control variables are *chSVI*, *BM*, *Ln(Size)*, *R_EA*, *# of News*, *IO*, *LnAF*, *MF*, *EV*, *EP*, and *Turnover*. *Ln(Size)*, *BM*, *IO*, *LnAF* and *Turnover* are measured as of the fiscal quarter end. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by calendar quarter and firm. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 7. Post Earnings Announcement Cumulative Abnormal Returns

	Random Walk Based Earnings Surprise (RW_SUE)		*Analyst Forecast Based Earnings Surprise (AF_SUE)*	
	(1)	***(2)***	***(3)***	***(4)***
R_SUE	0.053 ***	0.066 ***	0.055 ***	0.067 ***
	(0.014)	(0.023)	(0.015)	(0.021)
R_HQDisGD	-0.003	-0.003	-0.003	-0.003
	(0.004)	(0.007)	(0.004)	(0.007)
*R_SUE*R_HQDisGD*	-0.012 ***	-0.019 **	-0.017 **	-0.021 ***
	(0.005)	(0.009)	(0.007)	(0.009)
chSVI	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
*R_SUE*chSVI*	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.001)
# of News		-0.009		-0.008
		(0.008)		(0.008)
R_SUE# of News*		-0.008		-0.013
		(0.019)		(0.024)
Intercept	0.034	0.029	0.002	0.005
	(0.021)	(0.032)	(0.020)	(0.024)
Controls	Yes	Yes	Yes	Yes
*R_SUE*Controls*	Yes	Yes	Yes	Yes
State Fixed Effect	Yes	Yes	Yes	Yes
N	21597	13837	17842	10501
Adjusted R^2	2.05%	2.59%	2.12%	2.79%

This table presents the results of estimating ordinary least squares regressions. The dependent variable is cumulative abnormal return over a sixty trading day window starting two trading days after the earnings announcement date, *CAR[2,61]*. *HQDisGD* is measured in $year_{t-1}$. The full sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. The subsample with media coverage is from 2005 through 2008. The left-most columns display results using earnings surprise (SUE) values based on random walk expectations, while the right-most columns present results for SUE calculated using analyst-based expectations. The control variables are *chSVI, BM, Ln(Size), R_EA, # of News, IO, LnAF, MF, EV, EP*, and *Turnover*. *Ln(Size), BM, IO, LnAF* and *Turnover* are measured as of the fiscal quarter end. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by calendar quarter and firm. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 8. Changes in Geographic Dispersion and Market Reactions

Panel A. Random Walk Based (Absolute) Earnings Surprise

	AbSpreads [0,1]				**AbVol [0,1]**				**CAR [0,1]**				**CAR [2,61]**			
	Intercept: equation (3)				*Intercept: equation (4)*				*ERC: equation (5)*				*ERC: equation (6)*			
	Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$	
H	0.192	0.027	-0.165	***	0.584	1.716	1.132	***	0.050	0.081	0.031	***	0.062	0.024	-0.038	***
	(0.051)	(0.034)	(0.042)		(0.097)	(0.097)	(0.098)		(0.014)	(0.009)	(0.011)		(0.016)	(0.007)	(0.012)	
L	0.132	0.063	-0.069	*	0.628	0.703	0.075		0.051	0.055	0.004		0.057	0.049	-0.008	
	(0.045)	(0.033)	(0.038)		(0.176)	(0.164)	(0.171)		(0.015)	(0.005)	(0.009)		(0.015)	(0.005)	(0.009)	
H-L	0.060	-0.036	**-0.096**	**	-0.044	1.013***	**1.057**	***	-0.001	0.026***	**0.027**	***	0.005	-0.025***	**-0.030**	***
	(0.046)	(0.032)	**(0.039)**		(0.134)	(0.131)	**(0.134)**		(0.013)	(0.006)	**(0.009)**		(0.015)	(0.006)	**(0.011)**	

Panel B. Analyst Forecast Based (Absolute) Earnings Surprise

	AbSpreads [0,1]				**AbVol [0,1]**				**CAR [0,1]**				**CAR [2,61]**			
	Intercept: equation (3)				*Intercept: equation (4)*				*ERC: equation (5)*				*ERC: equation (6)*			
	Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$		Yr_{t-1}	Yr_{t+1}	$Yr_{t+1-t-1}$	
H	0.131	0.040	-0.091	***	0.623	1.736	1.113	***	0.078	0.121	0.043	***	0.057	0.021	-0.036	***
	(0.026)	(0.014)	(0.020)		(0.085)	(0.084)	(0.084)		(0.017)	(0.008)	(0.011)		(0.017)	(0.005)	(0.013)	
L	0.076	0.082	0.006		0.638	0.732	0.094		0.071	0.086	0.015		0.051	0.042	-0.009	
	(0.041)	(0.016)	(0.029)		(0.176)	(0.168)	(0.170)		(0.018)	(0.005)	(0.013)		(0.012)	(0.008)	(0.009)	
H-L	0.055	-0.042**	**-0.097**	***	-0.015	1.004***	**1.019**	***	0.007	0.035***	**0.028**	***	0.006	-0.021***	**-0.027**	**
	(0.034)	(0.017)	**(0.031)**		(0.132)	(0.125)	**(0.129)**		(0.016)	(0.006)	**(0.010)**		(0.014)	(0.006)	**(0.012)**	

This table presents the results of difference in difference analyses of 1) the changes in intercepts from estimating equations (3) and (4) but excluding the *R_HQDisGD* term, for *AbSpreads[0,1]* and *AbVol[0,1]* respectively, between high (H) change of *Ln(HQDisGD)* and low (L) change of *Ln(HQDisGD)*; 2) the changes in estimated coefficients of earnings surprise (*R_SUE)* from estimating equations (5) and (6) but excluding the *R_HQDisGD* and *R_SUE*R_HQDisGD* terms, for *CAR[0,1]* and *CAR[2,61]* respectively, between high (H) change of *Ln(HQDisGD)* and low (L) change of *Ln(HQDisGD)*. High (low) change of *Ln(HQDisGD)* is defined as top (bottom) tercile of change of *Ln(HQDisGD)* each year. The difference-in-difference coefficients and standard errors are shown in bold. The sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors, calculated using two-dimensional clustering by calendar quarter and firm. For the difference and difference-in-difference coefficients, ***, ** and * denotes statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.

Table 9. Propensity Score Matching Analysis

Panel A. Information Asymmetry

	Random Walk Based Absolute Earnings Surprise (RW_absSUE)						***Analyst Forecast Based Absolute Earnings Surprise (AF_absSUE)***					
AbSpread (0,1)	Treatment	Control	Difference	t-Statistic	p-Value		Treatment	Control	Difference	t-Statistic	p-Value	
Intercept of equation (3) without controls	0.041	0.068	-0.027	-2.20	0.028	**	0.027	0.063	-0.036	-2.72	0.007	***
Intercept of equation (3) with controls	0.033	0.073	-0.040	-3.05	0.002	***	0.068	0.148	-0.080	-3.22	0.001	***
AbVol (0,1)												
Intercept of equation (4) without controls	0.657	0.512	0.145	2.74	0.006	***	0.566	0.552	0.014	2.22	0.027	**
Intercept of equation (4) with controls	1.812	1.100	0.713	5.29	0.000	***	1.572	1.229	0.344	2.43	0.015	***

Panel B. Earnings Response Coefficients (ERC)

	Random Walk Based Earnings Surprise (RW_SUE)						***Analyst Forecast Based Earnings Surprise (AF_SUE)***					
CAR(0,1)	Treatment	Control	Difference	t-Statistic	p-Value		Treatment	Control	Difference	t-Statistic	p-Value	
ERC of equation (5) without controls	0.091	0.069	0.022	4.57	0.000	***	0.054	0.042	0.012	3.61	0.003	***
ERC of equation (5) with controls	0.052	0.025	0.028	2.61	0.009	***	0.068	0.032	0.036	2.79	0.005	***
CAR(2,61)												
ERC of equation (6) without controls	0.020	0.088	-0.068	-6.42	0.000	***	0.007	0.059	-0.052	-8.57	0.000	***
ERC of equation (6) with controls	0.079	0.105	-0.026	-1.98	0.048	**	0.008	0.050	-0.042	-3.18	0.002	***

This table presents the results of the difference in abnormal bid-ask spreads and abnormal trading volume (Panel A) and ERC during earnings announcement window and PEAD (Panel B) between the matched pairs of firms for different treatments. We model the conditional probability of having a certain level of the treatment conditional on economic characteristics in equation (2). The firms are matched according to the most similar conditional probability of treatment but the largest difference in the observed level of treatment. The left-most columns display results using (absolute) earnings surprise values based on random walk expectations, while the right-most columns present results for (absolute) earnings surprise calculated using analyst-based expectations. *, **, and *** denote statistical significance (two sided) at the 10%, 5%, and 1% levels, respectively.

Table 10. Path Analysis

	Random Walk Based Earnings Surprise (RW_SUE)		***Analyst Forcast Based Earnings Surprise (AF_SUE)***	
Direct Path	*CAR (0,1)*	*CAR (2,61)*	*CAR (0,1)*	*CAR (2,61)*
*R_SUE*R_HQDisGD -> CAR*	0.026 ***	-0.020 ***	0.025 ***	-0.019 **
	(0.006)	(0.006)	(0.007)	(0.007)
Indirect Path: Trading Volume				
R_HQDisGD -> Vol [0,1]	0.043 ***	0.043 ***	0.042 ***	0.042 ***
	(0.006)	(0.006)	(0.007)	(0.007)
*R_SUE*Vol[0,1] -> CAR*	0.209 ***	-0.138 ***	0.256 ***	-0.181 ***
	(0.031)	(0.031)	(0.040)	(0.043)
Total Trading Volume Path	0.009	-0.006	0.011	-0.008
Indirect Path: Investor Sophistication				
R_HQDisGD -> College_Ed	0.037 ***	0.037 ***	0.032 ***	0.032 ***
	(0.006)	(0.006)	(0.007)	(0.007)
*R_SUE*College_Ed -> CAR*	0.271 ***	-0.116 **	0.157 ***	-0.193 ***
	(0.038)	(0.036)	(0.040)	(0.043)
Total Investor Sophistication Path	0.010	-0.004	0.005	-0.006
Total Effect	0.045	-0.030	0.041	-0.033
Direct as a % of Total	58%	66%	61%	58%
Trading Volume as a % of Total	20%	20%	26%	23%
Investor Sophistication as a % of Total	22%	14%	12%	19%

This table reports the results from path analyses that examine the effect of geographic dispersion of Google search on return reactions to earnings news directly and through trading volume *(Vol[0,1])* or investor education (*College_Ed*). The equations include a regression of the outcome variable, *CAR[0,1]* or *CAR[2,61]*, on the source variable, *R_SUE*R_HQDisGD,* and mediating variables, *R_SUE*Vol[0,1]* and *R_SUE*College_Ed*(equation (7)) and regression of *Vol[0,1]* or *College_Edu* on *R_HQDisGD* (equation (8)). In equation (7), the control variables are *chSVI, BM, Ln(Size), R_EA, # of News, LnAF, MF, EV, EP*, and *IO.* In equation (8), the control variables are *chSVI, BM, Ln(Size), # of News, LnAF*, and *IO.* We present the standardized path coefficients. The sample consists of quarterly earnings announcements of firms listed on the major exchanges from 2005 through 2011. Variable definitions are listed in Appendix A. Numbers in parentheses are standard errors. ***, **, and * denote statistical significance at the 0.01, 0.05 and 0.10 levels, respectively, using 2-tailed tests.